

Sequence Match Listing for  
09/964824

for SETA ID NO:137

RESULT 2  
LOCUS AB011095/c  
DEFINITION Homo sapiens mRNA for KIAA0523 protein, partial cds.  
ACCESSION AB011095  
VERSION AB011095.1 GI:3043569  
KEYWORDS KIAA0523 protein.  
SOURCE Homo sapiens male brain cDNA to mRNA, clone\_1ib:pbluescriptII SK plus clone:HG1394.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (sites) Nagase,T., Ishikawa,K., Miyajima,N., Tanaka,A., Kotani,H., Nomura,N. and Ohara,O.  
TITLE Prediction of the coding sequences of unidentified human genes. IX. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro  
JOURNAL DNA Res. 5 (1), 31-39 (1998)  
MEDLINE 98290545  
AUTHORS Ohara,O., Nagase,T. and Ishikawa,K.  
REFERENCE 2 (bases 1 to 5172)  
TITLE Direct Submission  
JOURNAL Submitted (13-FEB-1998) Osamu Ohara, Kazusa DNA Research Institute, DNA Technology, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp, Tel:+81-438-52-3913, Fax:+81-438-52-3914)  
FEATURES  
source  
location/Qualifiers  
1..5172  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="HG1394"  
/sex="male"  
/tissue\_type="brain"  
/clone\_1ib="pbluescriptII SK plus"  
1..5172  
/gene="KIAA0523"  
c1..1407

/gene="KIAA0523"  
/codon\_start=1  
/product="KIAA0523 protein"  
/protein\_id="BA25449.1"  
/db\_xref="GI:3043570"  
/translation="ELKRLRRMFHHPNSDQPPALGPPAPPAIHSRGTYICFSD DGHETLKGAVFYDLRKMTVSHCODACAESESYAGAEABECYGRNLRAVSYGLE CNHECKGKSGVCGAVDRLSYRVDELQPSRRKRIATYIGCTRLPENITHAFPSLI QANVTGTCGSGFSOKRFPALILRGNECTICATPPRNLDADMSYGGDDPEKQRLA EYCEVYOTPVQDTRCTDRRLFPNKSRYVALISFPAGANTMAHLIENATGYTGSY FDCSTLYNKGFGKGDHMRKSRRTICVTHESGREIEMFDSAILIRPYSILAEPRR KCGHLSGTAADRMWKSKEMPDFNYSASHMSHLDLKYGRLLVYHBEILRSILVP TLRENVAFILNVSYSERLLCVENNKGSFRRRRRSRSHDEPTPEMDLINGYIRYVD QALRDHNMVTGLPREYVPR"  
BASE COUNT 1191 a 1340 c 1492 g 1149 t  
ORIGIN  
Query Match 100.0%; Score 263; DB 9; Length 5172;  
Best Local Similarity 100.0%; Pred. No. 2,4e-68;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AACACATTAACAGCAATTTTACGTCATATACAAAAAAGTCCAGAGTAAAGCCAAATC 60  
DB 5172 AACACATTAACAGCAATTTTACGTCATATACAAAAAAGTCCAGAGTAAAGCCAAATC 60  
QY 61 TCAGCAGAGCTTGATCTCTACTTAAACAAATTTACCAAGACCTGATCTCTTCTGCC 120  
DB 5112 TCAGCAGAGCTTGATCTCTACTTAAACAAATTTACCAAGACCTGATCTCTTCTGCC 120  
QY 121 TCTCAACTCTCCCTTCAGTGTGTCAGCTTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTC 180  
DB 5052 TCTCAACTCTCCCTTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTC 180  
QY 181 CAAGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTC 240  
DB 4992 CAAGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTCAGTGTGTC 240  
QY 241 TTAACAAGCTTCTCCAGCAT 263  
DB 4932 TTAACAAGCTTCTCCAGCAT 4910

RESULT 3  
LOCUS AC007962/c  
DEFINITION Homo sapiens chromosome 17 clone 2511\_J\_5 map 17, \*\*\* SEQUENCING IN PROGRESS \*\*\*  
ACCESSION AC007962  
VERSION AC007962.1 GI:5348397  
KEYWORDS HTG: HTGS\_PHASE1.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 172091)  
TITLE Homo sapiens chromosome 17, clone 2511\_J\_5  
JOURNAL Unpublished  
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baker,J., Baldwin,J., Baran,N., Beckley,R., Benn,J., Brown,A., Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearellano,K., Deyayre,E., Devon,K., Dewar,K., Donelan,L., Doyle,M., Ferreira,P., FitzHugh,M., Forrest,C., Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G., Hagos,B., Heath,A., Horton,L., Howland,J.C., Jones,C., Kain,L., Karats,A., Lehoczy,J., Lien,C., Locke,K., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J., Molla,M., Morris,M., Morrow,J., Mychaleckyj,J., Naylor,J., Nijhoff,M., O'Connor,T., O'Donnell,P., Pavlin,B., Peterson,K., Pollara,V., O'Connor,T., O'Donnell,P., Pavlin,B., Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A., Testaye,S., Tortuella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,

23 10:00:33 2003

us-09-964-824c-137.rge

Page 3

TITLE  
JOURNAL  
COMMENT

Wheeler, J., Wu, X., Wyman, D., Ye, W. J. and Zody, M.

Submitted (03-JUL-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
All repeats were identified using RepeatMasker: Smit, A.F.A. &  
Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>.

\* NOTE: This is a 'working draft' sequence. It currently consists of 25 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 2475: contig of 2475 bp in length  
2476 4789: contig of 2314 bp in length  
4790 7000: contig of 2211 bp in length  
7001 9027: contig of 2027 bp in length  
9028 11768: contig of 2741 bp in length  
11769 15071: contig of 3303 bp in length  
15072 19326: contig of 4255 bp in length  
19327 22239: contig of 2913 bp in length  
22240 26006: contig of 3767 bp in length  
26007 30364: contig of 4358 bp in length  
30365 34450: contig of 4086 bp in length  
34451 38286: contig of 3836 bp in length  
38287 42572: contig of 4286 bp in length  
42573 47699: contig of 5127 bp in length  
47700 54016: contig of 6317 bp in length  
54017 60547: contig of 6531 bp in length  
60548 68280: contig of 7733 bp in length  
68281 75214: contig of 6934 bp in length  
75215 80915: contig of 5701 bp in length  
80916 87222: contig of 6307 bp in length  
87223 94562: contig of 7340 bp in length  
94563 106577: contig of 12015 bp in length  
106578 119209: contig of 12632 bp in length  
119210 141857: contig of 22648 bp in length  
141858 172091: contig of 30234 bp in length.

FEATURES  
source

Location/Qualifiers

1. 172091

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="17"

/map="17"

/clone="2511\_J-5"

/clone.lib="CITB Human BAC"

BASE COUNT 44773 a 39840 c 42000 g 44968 t 510 others

ORIGIN

Query Match 100.0%; Score 263; DB 2; Length 172091;  
Best Local Similarity 100.0%; Pred. No. 3.9e-66; Mismatches 0; Gaps 0;  
Matches 263; Conservative 0; Indels 0; Gaps 0;

1 AACACATTAACAGATTATTTATAGTCATATATACAAAAAAGTCAGAGGTAAGGCCAATC 60  
DB 80393 AACACATTAACAGATTATTTATAGTCATATATACAAAAAAGTCAGAGGTAAGGCCAATC 80334  
QY 61 TCACAGACGCTTGATCTGTAATCTTAACAAATTCACAGACCTTGATCTTTCTGTC 120  
DB 80333 TCACAGACGCTTGATCTGTAATCTTAACAAATTCACAGACCTTGATCTTTCTGTC 80274  
QY 121 TCTCAACTCTCCCTTCAGTGTGTCAGCTTCAGCTGATTCCTGTCATGATCCAGAGGCC 180  
DB 80273 TCTCAACTCTCCCTTCAGTGTGTCAGCTTCAGCTGATTCCTGTCATGATCCAGAGGCC 80214  
QY 181 CAAGGTGTCATCATTAAGACCCAGAGAAATCTACTACTTTTTCACATTCACAGGGGAA 240  
DB 80213 CAAGGTGTCATCATTAAGACCCAGAGAAATCTACTACTTTTTCACATTCACAGGGGAA 80154  
QY 241 TTAACACAGCTTCTACCCAGCAI 263  
DB 80153 TTAACACAGCTTCTACCCAGCAI 80131

## ALIGNMENTS

RESULT 1

ABL67103

ID ABL67103 standard; DNA; 263 BP.

XX

AC ABL67103;

XX

DT 15-MAY-2002 (first entry)

XX

DE Thyroid cancer related gene sequence SEQ ID NO:5440.

XX

KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.

XX

OS Homo sapiens.

XX

PN WO200194629-A2.

XX

PD 13-DEC-2001.

XX

PF 30-MAY-2001; 2001WO-US10838.

XX

PR 05-JUN-2000; 2000US-209473P.

PR 05-JUN-2000; 2000US-209531P.

PR 18-SEP-2000; 2000US-233133P.

PR 18-SEP-2000; 2000US-233617P.

PR 20-SEP-2000; 2000US-234009P.

PR 20-SEP-2000; 2000US-234034P.

PR 20-SEP-2000; 2000US-234052P.

PR 22-SEP-2000; 2000US-234509P.

PR 22-SEP-2000; 2000US-234567P.

for SEQ ID NO. 137

Set ID No: 137 - cont -

PR	25-SEP-2000	2000US-234923P
PR	25-SEP-2000	2000US-234924P
PR	25-SEP-2000	2000US-235077P
PR	25-SEP-2000	2000US-235082P
PR	25-SEP-2000	2000US-235134P
PR	25-SEP-2000	2000US-235280P
PR	26-SEP-2000	2000US-235637P
PR	26-SEP-2000	2000US-235638P
PR	27-SEP-2000	2000US-235711P
PR	27-SEP-2000	2000US-235720P
PR	27-SEP-2000	2000US-235840P
PR	27-SEP-2000	2000US-235863P
PR	28-SEP-2000	2000US-236028P
PR	28-SEP-2000	2000US-236032P
PR	28-SEP-2000	2000US-236033P
PR	28-SEP-2000	2000US-236034P
PR	28-SEP-2000	2000US-236109P
PR	28-SEP-2000	2000US-236111P
PR	29-SEP-2000	2000US-236842P
PR	29-SEP-2000	2000US-236891P
PR	02-OCT-2000	2000US-237172P
PR	02-OCT-2000	2000US-237173P
PR	02-OCT-2000	2000US-237278P
PR	02-OCT-2000	2000US-237294P
PR	02-OCT-2000	2000US-237295P
PR	02-OCT-2000	2000US-237316P
PR	03-OCT-2000	2000US-237425P
PR	03-OCT-2000	2000US-237598P
PR	03-OCT-2000	2000US-237604P
PR	03-OCT-2000	2000US-237605P
PR	03-OCT-2000	2000US-237606P
PR	01-NOV-2000	2000US-244867P
PR	01-NOV-2000	2000US-245084P

Db	1	AAACATTAACAGCAATTTATTAGCTCATATACAAAAAAGGTCAGAGGTAAGGCCATC	60
Qy	61	TCACAGAAAGCTGTATCCCTGTACTAAACAAATTCACCAAGAGATGTATCTTTTCGCC	120
Db	61	TCACGAAAGCTGTATCCCTGTACTTTAAACAAATTCACCAAGAGATGTATCTTTTCGCC	120
Qy	121	TCCTCACTCTCCCTTCAGTGTGTCTACGTGATTCCTGGTCATGTATCCCAAGGCC	180
Db	121	TCCTCACTCTCCCTTCAGTGTGTCTACGTGATTCCTGGTCATGTATCCCAAGGCC	180
Qy	121	TCCTCACTCTCCCTTCAGTGTGTCTACGTGATTCCTGGTCATGTATCCCAAGGCC	180
Db	121	TCCTCACTCTCCCTTCAGTGTGTCTACGTGATTCCTGGTCATGTATCCCAAGGCC	180
Qy	181	CAAGTGTGTATCATTAAGACCCAGGAATACTACTCTTTTTCATCTCAACAGGGGAA	240
Db	181	CAAGTGTGTATCATTAAGACCCAGGAATACTACTCTTTTTCATCTCAACAGGGGAA	240
Qy	241	TTTAAACAGCTTCTACCCAGCAT	263
Db	241	TTTAAACAGCTTCTACCCAGCAT	263

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2003, 20:42:18 ; Search time 779.199 Seconds

(without alignments)  
5466.401 Million cell updates/sec

Title: US-09-964-824C-137

Perfect score: 263

Sequence: 1 aaacaataacagatctat.....aaacagctctaccagcat 263

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Capext 1.0

Searched: 16154066 seqs, 8097743376 residues 32308132

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estlun:\*  
5: em\_estlov:\*  
6: em\_estlpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_estl1:\*  
10: gb\_estl2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estlom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vtl:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	263	100.0	263	9	AA426220 zv84d01.s
2	263	100.0	300	9	AI915364
3	263	100.0	309	14	BM930947 UT-E-EJ0-
4	263	100.0	312	14	BM684449
5	263	100.0	348	14	Z39512
6	263	100.0	352	14	N63086

7	263	100.0	397	9	AI095108	AI095108 ga14h11.x
8	263	100.0	440	9	AI056703	AI056703 oys3e04.x
9	263	100.0	455	10	BE326640	BE326640 hys2f06.x
10	263	100.0	470	9	AI359305	AI359305 qy29c09.x
11	263	100.0	481	9	AI421275	AI421275 ft19g05.x
12	263	100.0	491	10	AM008726	AM008726 ws72c12.x
13	263	100.0	497	9	AI685123	AI685123 wc67g02.x
14	263	100.0	512	10	AM294233	AM294233 ut-H-B12-
15	263	100.0	522	14	BM676160	BM676160 ut-E-EJ0-
16	263	100.0	540	9	AA129740	AA129740 zpn1e08.s
17	263	100.0	541	9	AI628620	AI628620 ly77c03.x
18	263	100.0	547	10	BE504506	BE504506 h257d03.x
19	263	100.0	564	14	BM682221	BM682221 ut-E-EJ0-
20	263	100.0	588	10	AM290920	AM290920 ut-H-B12-
21	262	99.6	270	14	F02819	F02819 HSC1BH072.n
22	261.4	99.4	583	9	AA938912	AA938912 op74f09.s
23	261.4	99.4	583	9	AI918307	AI918307 we29c10.x
24	258.2	98.2	272	9	AI191041	AI191041 qe25a07.x
25	257.8	98.0	509	14	H29606	H29606 ym61d05.s1
26	256.8	97.6	317	14	T89670	T89670 yd99d03.s1
27	251	95.4	405	9	AA992195	AA992195 ot80e07.s
28	251	95.4	666	14	BM723364	BM723364 ut-E-EJ0-
29	247	93.9	447	14	H08230	H08230 y192g05.s1
30	246	93.5	432	14	H24358	H24358 ym56b01.s1
31	245	93.2	413	14	T91053	T91053 yd59c06.s1
32	243	92.4	262	9	AI686809	AI686809 tcs1904.x
33	240.6	91.5	548	17	AO701917	AO701917 HS-5396.B
34	236.4	89.9	374	10	AM136979	AM136979 ut-H-B11-
35	234.8	89.3	512	12	BG170361	BG170361 602323264
36	233.4	88.7	417	14	H66033	H66033 yr70b02.s1
37	228.4	86.8	420	14	H23292	H23292 ym57b08.s1
38	204	77.6	234	14	T31702	T31702 EST37299.Hu
39	186.8	71.0	761	12	BG108592	BG108592 602278574
40	178.6	67.9	392	14	H18630	H18630 ym45a08.s1
41	150.8	57.3	448	12	BF155474	BF155474 OY4-HM053
42	144	54.8	375	14	BM712977	BM712977 ut-E-EJ0-
43	139.2	52.9	194	9	AI609882	AI609882 t63e08.x
44	138.2	52.5	181	9	AI610939	AI610939 t60e06.x
45	113.4	43.1	655	12	BF944654	BF944654 PM0-NN117

ALIGNMENTS

RESULT 1  
AA426220 263 bp mRNA linear EST 16-OCT-1997  
DEFINITION zv84d01.s1 Soares\_total\_fetus\_ND2HF8\_9w Homo sapiens cDNA clone  
IMAGE:760321 3', mRNA sequence.  
AA426220  
AA426220.1 GI:2107753

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

human.  
Homo sapiens  
Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.,  
Schellenberg,K., Steptoe,K., Tan,F., Theising,B., White,T., Wyllie  
'T., Waterston,R. and Wilson,R.  
Washu-Merck EST Project 1997  
Unpublished (1997)  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu

TITLE  
JOURNAL  
COMMENT

This clone is available royalty-free through LINT; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -41ml3 fwd. RT from Amersham  
High quality sequence stop: 194.

FEATURES  
source  
location/Qualifiers  
1. 263  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="Soares\_NFL\_T.GBC.SI"  
/dev\_stage="8-9 weeks"  
/lab\_host="DH10B"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site 1: Not I. Site 2: Eco RI. 1st strand cDNA was prepared from mRNA obtained from pooled 8-9 week (total) fetus material with a Not I - 0190(dtt) primer [5' TGTACCAATCTGACAGCGAGCGCGCTTAATTTTCTTTTCTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT  
ORIGIN  
85 a 67 c 42 g 69 t

Query Match  
Best Local Similarity 100.0%; Score 263; DB 9; Length 263;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAACATTAACAGATTATTAGCTCATATTAACAAAAAGTCCAGAGTAAGCCATC 60  
DB 1 AAACATTAACAGATTATTAGCTCATATTAACAAAAAGTCCAGAGTAAGCCATC 60  
QY 61 TCAGCAGAGCTTGATCTGCTACTTAACATTTACCAAGAGCTGATCTTTTGGC 120  
DB 61 TCAGCAGAGCTTGATCTGCTACTTAACATTTACCAAGAGCTGATCTTTTGGC 120  
QY 121 TCTCAACTCTCCCTTAGTGTGTGTCAGCTTACGATTCCTGTCATGATGCCAAGGCC 180  
DB 121 TCTCAACTCTCCCTTAGTGTGTGTCAGCTTACGATTCCTGTCATGATGCCAAGGCC 180  
QY 121 TCTCAACTCTCCCTTAGTGTGTGTCAGCTTACGATTCCTGTCATGATGCCAAGGCC 180  
DB 121 TCTCAACTCTCCCTTAGTGTGTGTCAGCTTACGATTCCTGTCATGATGCCAAGGCC 180  
QY 181 CAAGGTGTCATCATTAAGACCCAGAAATCTACTACCTTTTACATTAACAGGGGAA 240  
DB 181 CAAGGTGTCATCATTAAGACCCAGAAATCTACTACCTTTTACATTAACAGGGGAA 240  
QY 241 TTTAAACAGCTTCTTACCCAGCAT 263  
DB 241 TTTAAACAGCTTCTTACCCAGCAT 263

RESULT 2  
A1915364 300 bp mRNA linear EST 17-DEC-1999  
LOCUS A1915364  
DEFINITION wd38f08.x1 Soares\_NFL\_T.GBC.SI Homo sapiens cDNA clone  
ACCESSION IMAGE:2330439.3 mRNA sequence.  
VERSION A1915364  
KEYWORDS A1915364.1 GI:5635219  
SOURCE EST.  
ORGANISM human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 300)  
AUTHORS NCI-GAP http://www.ncbi.nlm.nih.gov/ncigap.  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaps-r@mail.nih.gov  
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 459 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 272.  
location/Qualifiers  
1. 300  
/organism="Homo sapiens"

FEATURES  
source  
location/Qualifiers  
1. 263  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="Soares\_NFL\_T.GBC.SI"  
/dev\_stage="8-9 weeks"  
/lab\_host="DH10B"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site 1: Not I. Site 2: Eco RI. Equal amounts of plasmid DNA from three normalized libraries (fetal lung NbH19, testis NHT, and B-cell NCI-GAP.GCB1) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT  
ORIGIN  
91 a 75 c 46 g 88 t

Query Match  
Best Local Similarity 100.0%; Score 263; DB 9; Length 300;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAACATTAACAGATTATTAGCTCATATTAACAAAAAGTCCAGAGTAAGCCATC 60  
DB 1 AAACATTAACAGATTATTAGCTCATATTAACAAAAAGTCCAGAGTAAGCCATC 70  
QY 61 TCAGCAGAGCTTGATCTGCTACTTAACATTTACCAAGAGCTGATCTTTTGGC 120  
DB 71 TCAGCAGAGCTTGATCTGCTACTTAACATTTACCAAGAGCTGATCTTTTGGC 130  
QY 121 TCTCAACTCTCCCTTAGTGTGTGTCAGCTTACGATTCCTGTCATGATGCCAAGGCC 180  
DB 121 TCTCAACTCTCCCTTAGTGTGTGTCAGCTTACGATTCCTGTCATGATGCCAAGGCC 190  
QY 121 TCTCAACTCTCCCTTAGTGTGTGTCAGCTTACGATTCCTGTCATGATGCCAAGGCC 190  
DB 121 TCTCAACTCTCCCTTAGTGTGTGTCAGCTTACGATTCCTGTCATGATGCCAAGGCC 190  
QY 181 CAAGGTGTCATCATTAAGACCCAGAAATCTACTACCTTTTACATTAACAGGGGAA 240  
DB 191 CAAGGTGTCATCATTAAGACCCAGAAATCTACTACCTTTTACATTAACAGGGGAA 250  
QY 241 TTTAAACAGCTTCTTACCCAGCAT 263  
DB 251 TTTAAACAGCTTCTTACCCAGCAT 273

RESULT 3  
BM930947 309 bp mRNA linear EST 13-MAR-2002  
LOCUS BM930947/c  
DEFINITION U1-E-EU-aip-a-08-0-01.r1 U1-E-EU Homo sapiens cDNA clone  
ACCESSION BM930947 U1-E-EU-aip-a-08-0-01 5', mRNA sequence.  
VERSION BM930947.1 GI:19390120  
KEYWORDS BM930947  
SOURCE EST.  
ORGANISM human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 309)  
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
Normalization and subtraction: two approaches to facilitate gene discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
COMMENT Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: msoares@blue.weeg.uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
CDNA library preparation: Dr. M. Bento Soares, University of Iowa  
CDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).  
The following repetitive elements were found in this cDNA sequence: 276-305, >AT-rich#low\_complexity (matched complement)  
Seq primer: M13 REVERSE

FEATURES  
Source  
Location/Qualifiers

1..309  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="UI-E-EJ0-aip-a-08-0-UI"  
/clone\_lib="UI-E-EJ0"  
/tissue\_type="fetal eyes, lens, eye anterior segment, optic nerve, retina, Retina Foveal and Macular, RPE and Choroid"  
/dev\_stage="fetal and adult"  
/lab\_host="DH10B (Life Technologies) (TI phage resistant)"  
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site1: EcoR I; Site2: Not I; UI-E-EJ0 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are: fetal eyes, AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment, AATGCCCAT; optic nerve, CCATTAGTG; Retina, CCGCG; Retina Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

BASE COUNT  
ORIGIN  
97 a 45 c 76 g 91 t

Query Match  
Best Local Similarity 100.0%; Score 263; DB 14; Length 309;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACCAATTAACAGATTATTTAGCTCATTAACAAAAAAGTCCAGAGTAAGCCCAATC 60  
Db 285 AACCAATTAACAGATTATTTAGCTCATTAACAAAAAAGTCCAGAGTAAGCCCAATC 226  
QY 61 TCAAGCAAGGCTTATCTGTAATTAACATTTCCACCAAGACTTGATCTTTCTGCC 120  
Db 225 TCAAGCAAGGCTTATCTGTAATTAACATTTCCACCAAGACTTGATCTTTCTGCC 166  
QY 121 TCTCAACTCTCCCTTCAGTGGTGCATGCTGATCTGCTGATGATCCCAAGGCC 180  
Db 165 TCTCAACTCTCCCTTCAGTGGTGCATGCTGATCTGCTGATGATCCCAAGGCC 106  
QY 181 CAAGTGCTATCATTAAGACCCAGATTAATTAATTTTACATTCACCAAGGGGAA 240  
Db 105 CAAGTGCTATCATTAAGACCCAGATTAATTAATTTTACATTCACCAAGGGGAA 46  
QY 241 TTTAAACAGCTTACCCAGCAT 263  
Db 45 TTTAAACAGCTTACCCAGCAT 23

RESULT 4  
LOCUS BM684449

DEFINITION UI-E-EJ0-aip-a-08-0-UI.s1 UI-E-EJ0 Homo sapiens cDNA clone  
VERSION BM684449  
KEYWORDS BM684449.1 GI:18994345  
EST.

SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477

COMMENT  
Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: msoares@blue.weeg.uiowa.edu

Tissue Procurement: Dr. Gregg Hageman  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).  
The following repetitive elements were found in this cDNA sequence: 1-28, >AT-rich#low\_complexity (matched complement)  
Seq primer: M13 Forward  
POLYA-Yes.

FEATURES  
Source  
Location/Qualifiers

1..312  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="UI-E-EJ0-aip-a-08-0-UI"  
/clone\_lib="UI-E-EJ0"  
/tissue\_type="fetal eyes, lens, eye anterior segment, optic nerve, retina, Retina foveal and Macular, RPE and Choroid"  
/dev\_stage="fetal and adult"  
/lab\_host="DH10B (Life Technologies) (TI phage resistant)"  
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site1: EcoR I; Site2: Not I; UI-E-EJ0 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are: fetal eyes, AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment, AATGCCCAT; optic nerve, CCATTAGTG; Retina, CCGCG; Retina foveal and Macular, GTCC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).  
TAG\_LIB=UI-E-EJ0  
TAG\_TISSUE=Foveal and Macular Retina  
TAG\_SEQ=GTCC"

BASE COUNT  
ORIGIN  
91 a 79 c 45 g 97 t

Query Match  
Best Local Similarity 100.0%; Score 263; DB 14; Length 312;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACCAATTAACAGATTATTTAGCTCATTAACAAAAAAGTCCAGAGTAAGCCCAATC 60  
Db 19 AACCAATTAACAGATTATTTAGCTCATTAACAAAAAAGTCCAGAGTAAGCCCAATC 78  
QY 61 TCAAGCAAGGCTTATCTGTAATTAACATTTCCACCAAGACTTGATCTTTCTGCC 120  
Db 79 TCAAGCAAGGCTTATCTGTAATTAACATTTCCACCAAGACTTGATCTTTCTGCC 138  
QY 121 TCTCAACTCTCCCTTCAGTGGTGCATGCTGATCTGCTGATGATCCCAAGGCC 180  
Db 139 TCTCAACTCTCCCTTCAGTGGTGCATGCTGATCTGCTGATGATCCCAAGGCC 198

QY 181 CAAGTGTGTCATATAAGACCAGATACCTACTACCTTTTTCACATTCACAGGGGAA 240  
|||||  
Db 199 CAAGGTGTCATATAAGACCAGATACCTACTACTCTTTTTCACATTCACAGGGGAA 258  
|||||  
QY 241 TTTAAACAGCTTCTACCCAGCAT 263  
|||||  
Db 259 TTTAAACAGCTTCTACCCAGCAT 281  
|||||

RESULT 5  
239512 348 bp mRNA linear EST 31-OCT-1994  
LOCUS HSC1BH062 normalized infant brain cDNA Homo sapiens cDNA clone  
DEFINITION C-1bh06 3', mRNA sequence.  
ACCESSION 239512  
VERSION 239512.1 GI:562704  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 348)  
Aulfray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C., Devignes  
M.D., Duprat,S., Houigatier,R., Jumeau,M.N., Lamy,B., Lorenzo,F.,  
Mitchell,H., Mariage-Samson,R., Pietu,G., Pouliot,Y.,  
Sebastiani,Kabackhis,C. and Tessier,A.  
IMAGE: molecular integration of the analysis of the human genome  
and its expression  
C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)  
95277534  
CONTACT: Genethon  
GenexPress-Genethon  
Genethon Centre de recherche sur le Genome Humain  
1, rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE  
Tel: 33160778698  
Fax: 33160778698  
Email: genexpress@genethon.fr  
Single read: 25 T removed at sequence 5'end  
GenexPress\_library\_id: C; GenexPress\_sequence\_id: alc-1bh06  
Seq primer: (21)M13-universal.  
Location/Qualifiers  
1. 348  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="C-1bh06"  
/clone\_lib="normalized infant brain cDNA"  
/sex="Female"  
/tissue\_type="total brain"  
/dev\_stage="3 months old"  
/note="Organ: brain; Vector: lambda BA; Site\_1: HindIII;  
Site\_2: NotI; sex:Female; dev\_stage=3 months old;  
isolate=muscular atrophy patient; tissue\_type=total brain  
; total mRNA was oligo-(dT) primed and directionally  
cloned 5' -> 3' into the HindIII -> NotI sites of the  
lambda BA vector. Clone library from B.Souares, Psychiatry  
Dept. Columbia University, USA. Normalization\_method:  
Bento Soares. P.N.A.S in press"

BASE COUNT 109 a 91 c 54 g 93 t 1 others

ORIGIN

Query Match 100.0%; Score 263; DB 14; Length 348;  
Best Local Similarity 100.0%; Pred. No. 4; 6e-61;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAACATTAACAGATTTATTAGTCATATTAACAAAAAAGTCACAGAGTAAGCCCAATC 60  
|||||  
Db 1 AAACATTAACAGATTTATTAGTCATATTAACAAAAAAGTCACAGAGTAAGCCCAATC 60  
|||||

QY 61 TCAGCAAGGCTTGATCTGCTACTTAACAATTTCCACAGGACTGATCTCTTTTGCC 120  
|||||  
Db 61 TCAGCAAGGCTTGATCTGCTACTTAACAATTTCCACAGGACTGATCTCTTTTGCC 120  
|||||

QY 121 TTCACACTGCCCTTCAGTGTGTGTCAGCTTCAGCGATTCCTGTCATATGCCAAGGCC 180  
|||||

Db 121 TTCACACTGCCCTTCAGTGTGTGTCAGCTTCAGCGATTCCTGTCATATGCCAAGGCC 180  
|||||

QY 181 CAAGTGTGTCATATAAGACCAGATACCTACTACTCTTTTTCACATTCACAGGGGAA 240  
|||||  
Db 181 CAAGGTGTCATATAAGACCAGATACCTACTACTCTTTTTCACATTCACAGGGGAA 240  
|||||

QY 241 TTTAAACAGCTTCTACCCAGCAT 263  
|||||  
Db 241 TTTAAACAGCTTCTACCCAGCAT 263  
|||||

RESULT 6  
N63086 352 bp mRNA linear EST 30-JAN-1997  
LOCUS yz32e07.s1 Soares,multiple sclerosis\_2NBHSP Homo sapiens cDNA  
DEFINITION clone IMAGE:284772 3', mRNA sequence.  
ACCESSION N63086  
VERSION N63086.1 GI:1210915  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 352)  
Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,  
Chisoe,S., Dietrich,N., Dubuque,T., Favelllo,A., Gish,W., Hawkins  
M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore  
B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,  
Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J., Trevisan,E.,  
Underwood,K., Woldmann,P., Waterson,R., Wilson,R. and Merrin,M.  
Generation and analysis of 280,000 human expressed sequence tags  
Genome Res. 6 (9), 807-828 (1996)  
97044478  
CONTACT: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LNLN; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 724 Std Error: 0.00  
Seq primer: m13 -40 forward  
High quality sequence stop: 285.  
Location/Qualifiers  
1. 352  
/organism="Homo sapiens"  
/db\_xref="GDB:3903300"  
/db\_xref="taxon:9606"  
/clone\_lib="IMAGE:284772"  
/clone\_lib="Soares-multiple-sclerosis\_2NBHSP"  
/sex="male"  
/tissue\_type="multiple sclerosis lesions"  
/dev\_stage="Age 46"  
/lab\_host="DH10B (ampicillin resistant)"  
/note="Vector: p773D (Pharmacia) with a modified  
polylinker V\_TYPE: phagemid; Site\_1: Not I; Site\_2: Eco RI  
; 1st strand cDNA was primed with a Not I - oligo(dT)  
primer [5',  
TCTTACCATCTGCACTGGAGCGGCCGATTTTTTTTTTTTTTTT 3'],  
double-stranded cDNA was size selected, ligated to Eco RI  
adapters (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of a modified p773 vector  
(Pharmacia). Library went through one round of  
normalization to a Cot = 5. Library constructed by Bento  
Soares and M.Fatima Bonaldo. RNA from 4 multiple sclerosis  
lesions from one patient was kindly provided by Dr. Kevin  
G. Becker (NINDS/NIH)."

BASE COUNT 112 a 92 c 55 g 93 t

ORIGIN

Query Match 100.0%; Score 263; DB 14; Length 352;

Best Local Similarity 100.0%; Pred. No. 4.6e-61;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACAAATTAACAGATTATTTAGTCTATATTAACAAAAAGTCCAGAGTAAGCCAAATC 60  
1 AAACAAATTAACAGATTATTTAGTCTATATTAACAAAAAGTCCAGAGTAAGCCAAATC 60  
Db 1 AAACAAATTAACAGATTATTTAGTCTATATTAACAAAAAGTCCAGAGTAAGCCAAATC 60

QY 61 TCAAGCAGAGCTTGATCTCTTACTTAAACAATTTCCACCAAGAGCTGATCTTTTGCC 120  
61 TCAAGCAGAGCTTGATCTCTTACTTAAACAATTTCCACCAAGAGCTGATCTTTTGCC 120  
Db 61 TCAAGCAGAGCTTGATCTCTTACTTAAACAATTTCCACCAAGAGCTGATCTTTTGCC 120

QY 121 TCCCAACTGCCCTTCAGTGGTGTGACCTTCAGCTGATGCTGATGATCCCAAGGCC 180  
121 TCCCAACTGCCCTTCAGTGGTGTGACCTTCAGCTGATGCTGATGATCCCAAGGCC 180  
Db 121 TCCCAACTGCCCTTCAGTGGTGTGACCTTCAGCTGATGCTGATGATCCCAAGGCC 180

QY 181 CAAGGTGGTCATCATTAAGAAGCCAGAAATCTACTACTCTTTTCACATTCACAGGGAA 240  
181 CAAGGTGGTCATCATTAAGAAGCCAGAAATCTACTACTCTTTTCACATTCACAGGGAA 240  
Db 181 CAAGGTGGTCATCATTAAGAAGCCAGAAATCTACTACTCTTTTCACATTCACAGGGAA 240

QY 241 TTAACACAGCTTCTACCCAGCAT 263  
241 TTAACACAGCTTCTACCCAGCAT 263  
Db 241 TTAACACAGCTTCTACCCAGCAT 263

RESULT 7  
LOCUS A1095108 397 bp mRNA linear EST 01-OCT-1998  
DEFINITION ga14h11.x1 NCI\_CGAP\_Brn23 Homo sapiens cDNA clone IMAGE:1686789 3',  
mRNA sequence.  
ACCESSION A1095108  
VERSION A1095108.1 GI:3434084  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 397)  
NCI/NINDS-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute / National Institute of Neurological  
Disorders and Stroke, Brain Tumor Genome Anatomy Project  
(CGAP/BRGAP), Tumor Gene Index  
Unpublished (1998)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,  
Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
Bonaldo, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
Insert length: 724 Std Error: 0.00  
Seq primer: -40ml3 fwd. RT from Amersham  
High quality sequence stop: 396.  
Location/Qualifiers  
1. 397  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1686789"  
/clone\_id="NCI\_CGAP\_Brn23"  
/tissue\_type="glioblastoma (pooled)"  
/lab\_host="DH108"  
/note="Organ: brain; Vector: pT773D-Pac (Pharmacia) with a  
modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st  
strand cDNA was primed with a Not I - oligo(dT) primer [5'  
TGTACCAATCTGAGTGGAGCGCCGCAATCTTTTGTGTGTGTGTGTGTGTGTGTGTGT  
T 3']; double-stranded cDNA was ligated to Eco RI  
adaptors (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of the modified pT773 vector.  
Library is normalized, and was constructed by Bento  
Soares and M. Fatima Bonaldo."

BASE COUNT 121 a 102 c 68 g 106 t  
ORIGIN

Query Match 100.0%; Score 263; DB 9; Length 397;  
Best Local Similarity 100.0%; Pred. No. 4.7e-61;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAACAAATTAACAGATTATTTAGTCTATATTAACAAAAAGTCCAGAGTAAGCCAAATC 60  
1 AAACAAATTAACAGATTATTTAGTCTATATTAACAAAAAGTCCAGAGTAAGCCAAATC 60  
Db 2 AAACAAATTAACAGATTATTTAGTCTATATTAACAAAAAGTCCAGAGTAAGCCAAATC 61

QY 61 TCAAGCAGAGCTTGATCTCTTACTTAAACAATTTCCACCAAGAGCTGATCTTTTGCC 120  
61 TCAAGCAGAGCTTGATCTCTTACTTAAACAATTTCCACCAAGAGCTGATCTTTTGCC 120  
Db 61 TCAAGCAGAGCTTGATCTCTTACTTAAACAATTTCCACCAAGAGCTGATCTTTTGCC 120

QY 121 TCCCAACTGCCCTTCAGTGGTGTGACCTTCAGCTGATGCTGATGATCCCAAGGCC 180  
121 TCCCAACTGCCCTTCAGTGGTGTGACCTTCAGCTGATGCTGATGATCCCAAGGCC 180  
Db 121 TCCCAACTGCCCTTCAGTGGTGTGACCTTCAGCTGATGCTGATGATCCCAAGGCC 180

QY 181 CAAGGTGGTCATCATTAAGAAGCCAGAAATCTACTACTCTTTTCACATTCACAGGGAA 240  
181 CAAGGTGGTCATCATTAAGAAGCCAGAAATCTACTACTCTTTTCACATTCACAGGGAA 240  
Db 181 CAAGGTGGTCATCATTAAGAAGCCAGAAATCTACTACTCTTTTCACATTCACAGGGAA 241

QY 241 TTAACACAGCTTCTACCCAGCAT 263  
241 TTAACACAGCTTCTACCCAGCAT 263  
Db 241 TTAACACAGCTTCTACCCAGCAT 264

RESULT 8  
LOCUS A1056703 440 bp mRNA linear EST 20-JUL-1998  
DEFINITION oy53e04.x1 NCI\_CGAP\_Brn23 Homo sapiens cDNA clone IMAGE:1669566 3',  
mRNA sequence.  
ACCESSION A1056703  
VERSION A1056703.1 GI:3330569  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 440)  
NCI/NINDS-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute / National Institute of Neurological  
Disorders and Stroke, Brain Tumor Genome Anatomy Project  
(CGAP/BRGAP), Tumor Gene Index  
Unpublished (1998)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,  
Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
Bonaldo, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
Seq primer: -40ml3 fwd. RT from Amersham.  
Location/Qualifiers  
1. 440  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1669566"  
/clone\_id="NCI\_CGAP\_Brn23"  
/tissue\_type="glioblastoma (pooled)"  
/lab\_host="DH108"  
/note="Organ: brain; Vector: pT773D-Pac (Pharmacia) with a  
modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st  
strand cDNA was primed with a Not I - oligo(dT) primer [5'  
TGTACCAATCTGAGTGGAGCGCCGCAATCTTTTGTGTGTGTGTGTGTGTGTGTGTGT  
T 3']; double-stranded cDNA was ligated to Eco RI  
adaptors (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of the modified pT773 vector.

Library is normalized, and was constructed by Bento Soares and M. Fatima Bonaldo.  
BASE COUNT 134 a 110 c 80 g 116 t  
ORIGIN

Query Match 100.0%; Score 263; DB 9; Length 440;  
Best Local Similarity 100.0%; Pred. No. 4,9e-61;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACCAATTAACAGATTATAGCTCATATACAAAAAAGTCAGAGTAAGCCATC 60  
DB 2 AACCAATTAACAGATTATAGCTCATATACAAAAAAGTCAGAGTAAGCCATC 61  
QY 61 TCAAGCAGGCTTGATCTCTACTTAAACAATTTCACCAAGAGCTGATCTTTGCC 120  
DB 62 TCAAGCAGGCTTGATCTCTACTTAAACAATTTCACCAAGAGCTGATCTTTGCC 121  
QY 121 TCTCACTCTCCCTTCAGTGTGTCAGCTTCAGTGTGTCATGATCCCAAGGCC 180  
DB 122 TCTCACTCTCCCTTCAGTGTGTCAGCTTCAGTGTGTCATGATCCCAAGGCC 181  
QY 181 CAAGGTGTCATATAAGACCAGAAATCTACTACTCTTTTCACATTCACAGGGAA 240  
DB 182 CAAGGTGTCATATAAGACCAGAAATCTACTACTCTTTTCACATTCACAGGGAA 241  
QY 241 TTAACAAGCTTCTACCCAGCAT 263  
DB 242 TTAACAAGCTTCTACCCAGCAT 264

RESULT 9  
BE326640 455 bp mRNA linear EST 14-JUL-2000  
LOCUS br62f06.x1 NCL\_CGAP\_Kid11 Homo sapiens cDNA clone IMAGE:3133091 3',  
DEFINITION mRNA sequence.  
ACCESSION BE326640  
VERSION BE326640.1 GI:9200416  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
COMMENT Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LINL, send email to:  
infoimage.lnl.gov

Seq primer: -40UP from Gibco.

FEATURES  
source  
1. 455  
Location/Qualifiers

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3133091"  
/clone\_lib="NCI\_CGAP\_Kid11"  
/lab\_host="DH10B"  
/note="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with  
a modified polylinker; Site:1: Not I; Site:2: Eco RI;  
Plasmid DNA from the normalized library NCI\_CGAP\_Kid3 was  
prepared, and ss circles were made in vitro. Following HAP  
purification, this DNA was used as tracer in a subtractive  
hybridization reaction. The driver was PCR-amplified cDNAs  
from a pool of 5,000 clones made from the same library  
(cloneids 132376-1323911, 1456007-1456775, and

1500552-1502855). Subtraction by Bento Soares and M.  
Fatima Bonaldo.  
BASE COUNT 140 a 111 c 87 g 117 t  
ORIGIN

Query Match 100.0%; Score 263; DB 10; Length 455;  
Best Local Similarity 100.0%; Pred. No. 4,9e-61;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACCAATTAACAGATTATAGCTCATATACAAAAAAGTCAGAGTAAGCCATC 60  
DB 1 AACCAATTAACAGATTATAGCTCATATACAAAAAAGTCAGAGTAAGCCATC 60  
QY 61 TCAAGCAGGCTTGATCTCTACTTAAACAATTTCACCAAGAGCTGATCTTTGCC 120  
DB 62 TCAAGCAGGCTTGATCTCTACTTAAACAATTTCACCAAGAGCTGATCTTTGCC 120  
QY 121 TCTCACTCTCCCTTCAGTGTGTCAGCTTCAGTGTGTCATGATCCCAAGGCC 180  
DB 122 TCTCACTCTCCCTTCAGTGTGTCAGCTTCAGTGTGTCATGATCCCAAGGCC 180  
QY 181 CAAGGTGTCATATAAGACCAGAAATCTACTACTCTTTTCACATTCACAGGGAA 240  
DB 182 CAAGGTGTCATATAAGACCAGAAATCTACTACTCTTTTCACATTCACAGGGAA 240  
QY 241 TTAACAAGCTTCTACCCAGCAT 263  
DB 242 TTAACAAGCTTCTACCCAGCAT 263

RESULT 10  
A1359305 470 bp mRNA linear EST 15-FEB-1999  
LOCUS gY29c09.x1 NCI\_CGAP\_Brn23 Homo sapiens cDNA clone IMAGE:2013424 3',  
DEFINITION mRNA sequence.  
ACCESSION A1359305  
VERSION A1359305.1 GI:4110926  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
TITLE NCI/MINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
JOURNAL National Cancer Institute / National Institute of Neurological  
Disorders and Stroke, Brain Tumor Genome Anatomy Project  
(CGAP/BRGAP), Tumor Gene Index  
Unpublished (1998)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,  
Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
Bonaldo, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LINL at:  
www.bio.lnl.gov/brp/image/image.html  
Insert Length: 730 Std Error: 0.00  
Seq primer: -40UP from Gibco.

FEATURES  
source  
1. 470  
Location/Qualifiers

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2013424"  
/clone\_lib="NCI\_CGAP\_Brn23"  
/tissue\_type="glioblastoma (pooled)"  
/lab\_host="DH10B"  
/note="Organ: Brain; Vector: pT7T3D-Pac (Pharmacia) with a  
modified polylinker; Site:1: Not I; Site:2: Eco RI; 1st  
strand cDNA was primed with a Not I - oligo(dT) primer [5'  
GTGTACCAATCTGAACTGGAGCGCCGCAATATCTTTTTTTTTTTTTTTTTTTT

T 3']: double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library is normalized, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 145 a 111 c 94 g 120 t  
ORIGIN

Query Match 100.0%; Score 263; DB 9; Length 470;  
Best Local Similarity 100.0%; Pred. No. 4.9e-61;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAACAATTAACAGATTATTAGCTCATATACAAAAAAGTCCAGAGTAAGCCATC 60  
Db 1 AAACAATTAACAGATTATTAGCTCATATACAAAAAAGTCCAGAGTAAGCCATC 60  
QY 61 TCAAGCAAGGCTTATCTCTGACTTAAACAATTTCACCAAGAGTGTGATCTCTTCC 120  
Db 61 TCAAGCAAGGCTTATCTCTGACTTAAACAATTTCACCAAGAGTGTGATCTCTTCC 120  
QY 121 TCTCACTCTCCCTTCAGTGGTGTACAGCTTCAGTGATTCCTGTGATGATCCAGGCC 180  
Db 121 TCTCACTCTCCCTTCAGTGGTGTACAGCTTCAGTGATTCCTGTGATGATCCAGGCC 180  
QY 181 CAAGGTGGTATCTAAAGACCCAGAACTACTACTCTTTTTCACATTCAACAGGGGAA 240  
Db 181 CAAGGTGGTATCTAAAGACCCAGAACTACTACTCTTTTTCACATTCAACAGGGGAA 240  
QY 241 TTAACAAGCTTCTACCCAGCAT 263  
Db 241 TTAACAAGCTTCTACCCAGCAT 263

RESULT 11 481 bp mRNA linear EST 28-MAR-1999  
A1421276  
LOCUS t119495.x1 NCI\_CGAP\_Brn23 Homo sapiens cDNA clone IMAGE:2096696 3',  
DEFINITION mRNA sequence.  
ACCESSION A1421276  
VERSION A1421276.1 GI:4267207  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 481)  
NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute / National Institute of Neurological  
Disorders and Stroke, Brain Tumor Genome Anatomy Project  
(CGAP/BrGAP), Tumor Gene Index  
Unpublished (1998)

JOURNAL  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,  
Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/dbp/image/image.html  
Insert length: 745 Std Error: 0.00  
Seq primer: -40UP from GIBCO  
High quality sequence stop: 451.  
Location/Qualifiers

## FEATURES

SOURCE

1..481  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2096696"  
/clone\_lib="NCI\_CGAP\_Brn23"  
/tissue\_type="glioblastoma (pooled)"  
/lab\_host="DH10B"

/note="Organ: brain; Vector: p773D-Pac (Pharmacia) with a  
modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st  
strand cDNA was primed with a Not I - oligo(dT) primer [5',  
TGTACCAATCTGTAAGGAGCGGCGGCAATATCTTTTCTTTTCTTTT  
T 3']: double-stranded cDNA was ligated to Eco RI  
adaptors (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of the modified p773 vector.  
Library is normalized, and was constructed by Bento  
Soares and M. Fatima Bonaldo."

BASE COUNT 149 a 114 c 96 g 122 t  
ORIGIN

Query Match 100.0%; Score 263; DB 9; Length 481;  
Best Local Similarity 100.0%; Pred. No. 5e-61;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAACAATTAACAGATTATTAGCTCATATACAAAAAAGTCCAGAGTAAGCCATC 60  
Db 1 AAACAATTAACAGATTATTAGCTCATATACAAAAAAGTCCAGAGTAAGCCATC 60  
QY 61 TCAAGCAAGGCTTATCTCTGACTTAAACAATTTCACCAAGAGTGTGATCTCTTCC 120  
Db 61 TCAAGCAAGGCTTATCTCTGACTTAAACAATTTCACCAAGAGTGTGATCTCTTCC 120  
QY 121 TCTCACTCTCCCTTCAGTGGTGTACAGCTTCAGTGATTCCTGTGATGATCCAGGCC 180  
Db 121 TCTCACTCTCCCTTCAGTGGTGTACAGCTTCAGTGATTCCTGTGATGATCCAGGCC 180  
QY 181 CAAGGTGGTATCTAAAGACCCAGAACTACTACTCTTTTTCACATTCAACAGGGGAA 240  
Db 181 CAAGGTGGTATCTAAAGACCCAGAACTACTACTCTTTTTCACATTCAACAGGGGAA 240  
QY 241 TTAACAAGCTTCTACCCAGCAT 263  
Db 241 TTAACAAGCTTCTACCCAGCAT 263

RESULT 12 491 bp mRNA linear EST 08-MAR-2000  
AW008726  
LOCUS w572c12.x1 NCI\_CGAP\_Brn23 Homo sapiens cDNA clone IMAGE:2502742 3',  
DEFINITION mRNA sequence.  
ACCESSION AW008726  
VERSION AW008726.1 GI:5857504  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 491)  
NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute / National Institute of Neurological  
Disorders and Stroke, Brain Tumor Genome Anatomy Project  
(CGAP/BrGAP), Tumor Gene Index  
Unpublished (1998)

JOURNAL  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,  
Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/dbp/image/image.html  
Insert length: 679 Std Error: 0.00  
Seq primer: -40UP from GIBCO  
High quality sequence stop: 465.  
Location/Qualifiers

## FEATURES

SOURCE

1..491  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

```
/clone_1lb="NCI_CGAP_Pan1"
/tissue_type="glioiblastoma (pooled)"
/lab_host="DH10B"
/note="Organ: brain; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTCACCAATCGACGTGAGCGGCGGCACTTCTTTTCTTTTCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
Library is normalized, and was constructed by Bento
Soares and M.Farima Bernaldo."
```

BASE COUNT 153 a 118 c 96 g 123 t 1 others

ORIGIN

Query Match 100.0%; Score 263; DB 10; Length 491;  
Best Local Similarity 100.0%; Pred. No. 5e-61;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACCATTAACAGATTATTTAGCTCATATACAAAAAAGTCCAGAGGTAAAGCCATC 60  
Db 1 AACCATTAACAGATTATTTAGCTCATATACAAAAAAGTCCAGAGGTAAAGCCATC 60  
QY 61 TCAGCAGAGGCTTGATCTGTACTTAACAAATTTACACAGAGCTGATCTTTGCGC 120  
Db 61 TCAGCAGAGGCTTGATCTGTACTTAACAAATTTACACAGAGCTGATCTTTGCGC 120  
QY 121 TCTCAACTCTCCCTTCAAGTGTGTCAAGCTTCAAGTATCTTGTCATGATCCCAAGGCC 180  
Db 121 TCTCAACTCTCCCTTCAAGTGTGTCAAGCTTCAAGTATCTTGTCATGATCCCAAGGCC 180  
QY 121 TCTCAACTCTCCCTTCAAGTGTGTCAAGCTTCAAGTATCTTGTCATGATCCCAAGGCC 180  
Db 121 TCTCAACTCTCCCTTCAAGTGTGTCAAGCTTCAAGTATCTTGTCATGATCCCAAGGCC 180  
QY 181 CAAGGTGGTCATCATATAAGACCCAGGATATCTACTTCTTTCACATTCACAGGGAA 240  
Db 181 CAAGGTGGTCATCATATAAGACCCAGGATATCTACTTCTTTCACATTCACAGGGAA 240  
QY 241 TTTAAACAGCTTCTACCCAGCAT 263  
Db 241 TTTAAACAGCTTCTACCCAGCAT 263

RESULT 13  
A1685123 497 bp mRNA linear EST 17-DEC-1999  
LOCUS wc67602.x1 NCI\_CGAP\_Pan1 Homo sapiens cDNA clone IMAGE:2323730 3',  
DEFINITION mRNA sequence.  
ACCESSION A1685123  
VERSION A1685123.1 GI:4896417  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 497)  
REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
TITLE Unpublished (1997)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgabbs-remail.nih.gov  
Life Technologies catalog #: 11548-013  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www.bio.llnl.gov/dbp/image/image.html  
Insert Length: 766 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 412.  
Location/Qualifiers  
1..497  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2323730"

```
/clone_1lb="NCI_CGAP_Pan1"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/note="Organ: pancreas; Vector: pCMV-Sport6; Site_1: SalI;  
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.72 Kb. Life Technologies catalog #:  
11548-013"
```

BASE COUNT 154 a 120 c 97 g 126 t

ORIGIN

Query Match 100.0%; Score 263; DB 9; Length 497;  
Best Local Similarity 100.0%; Pred. No. 5e-61;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACCATTAACAGATTATTTAGCTCATATACAAAAAAGTCCAGAGGTAAAGCCATC 60  
Db 2 AACCATTAACAGATTATTTAGCTCATATACAAAAAAGTCCAGAGGTAAAGCCATC 61  
QY 61 TCAGCAGAGGCTTGATCTGTACTTAACAAATTTACACAGAGCTGATCTTTGCGC 120  
Db 62 TCAGCAGAGGCTTGATCTGTACTTAACAAATTTACACAGAGCTGATCTTTGCGC 121  
QY 121 TCTCAACTCTCCCTTCAAGTGTGTCAAGCTTCAAGTATCTTGTCATGATCCCAAGGCC 180  
Db 122 TCTCAACTCTCCCTTCAAGTGTGTCAAGCTTCAAGTATCTTGTCATGATCCCAAGGCC 181  
QY 181 CAAGGTGGTCATCATATAAGACCCAGGATATCTACTTCTTTCACATTCACAGGGAA 240  
Db 182 CAAGGTGGTCATCATATAAGACCCAGGATATCTACTTCTTTCACATTCACAGGGAA 241  
QY 241 TTTAAACAGCTTCTACCCAGCAT 263  
Db 242 TTTAAACAGCTTCTACCCAGCAT 264

RESULT 14  
AM294233 512 bp mRNA linear EST 16-JAN-2000  
LOCUS UT-H-B12-abc-f-12-0-UT.s1 NCI\_CGAP\_Sub4 Homo sapiens cDNA clone  
DEFINITION IMAGE:2726542 3', mRNA sequence.  
ACCESSION AM294233  
VERSION AM294233.1 GI:6700869  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 512)  
REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
TITLE Unpublished (1997)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgabbs-remail.nih.gov  
The sequence contained an oligo-dT track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A  
tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:  
NCI-CGAP clone distribution information can be found through the  
I.M.A.G.E. Consortium/LLNL at:  
www.bio.llnl.gov/dbp/image/image.html The following repetitive  
elements were found in this cDNA sequence: 1-27,  
>AT-rich#Low-complexity  
Seq primer: M13 Forward  
POLYA=Yes.  
Location/Qualifiers  
1..512  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2726542"  
/clone\_1lb="NCI\_CGAP\_Sub4"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker, Site\_1: Not I; Site\_2: Eco RI; The NCI\_CGAP\_Sub4 library is a subtracted library derived from the NCI\_CGAP\_Sub2 library which is a subtracted library derived from the NCI\_CGAP\_Sub1 library, which is a subtracted library derived from Bi. Bi constitutes a mixture of 2i normalized or subtracted NCI\_CGAP libraries: NCI\_CGAP\_Co4, NCI\_CGAP\_Pr22, NCI\_CGAP\_Pr28, NCI\_CGAP\_Co10, NCI\_CGAP\_Co16, NCI\_CGAP\_Kid5, NCI\_CGAP\_Kid12, NCI\_CGAP\_Kid3, NCI\_CGAP\_Co8, NCI\_CGAP\_Pr28, NCI\_CGAP\_Le12, NCI\_CGAP\_Brn23, NCI\_CGAP\_Lu5, NCI\_CGAP\_C1L1, NCI\_CGAP\_Lu24, NCI\_CGAP\_Lu19, NCI\_CGAP\_Co4, NCI\_CGAP\_Go6, NCI\_CGAP\_Brn25. These 2i libraries were pooled and a single-stranded DNA preparation of the resulting mixture was used as a tracer in a subtractive hybridization with a driver whose composition is detailed below:

NCI\_CGAP\_Kid3 pool 1: LLM 3334-3337, 3682-3683, 3798-3803 (IMAGE cloneids 1322376-1323811, 1456008-1456775, 1500552-1502853) NCI\_CGAP\_Kid5 pool 1: LLM 3338-3342, 3722-3725, 3776-3778 (IMAGE cloneids 1333912-1325831, 1471368-1472903, 1492104-1493255) NCI\_CGAP\_Lu5 pool 1: LLM 3575-3582, 3851-3854 (IMAGE cloneids 1414920-1417991, 1520904-1522439) NCI\_CGAP\_Co4 pool 1: LLM 3164-3167, 3716-3720, 3733-3735 (IMAGE cloneids 1257096-1258631, 1469064-1470993, 1475592-1476743) NCI\_CGAP\_Pr22 pool 1: LLM 2457-2459, 2758-2759, 3062-3068 (IMAGE cloneids 985608-986759, 1101192-1101959, 1217928-1220615) NCI\_CGAP\_Co10 pool 1: LLM 2644-2653, 2871-2872 (IMAGE cloneids 1057416-1061255, 1144584-1145351)

) Subtraction was performed as previously described [Bonaldo, Lennon & Soares (1996): Normalization and Subtraction: Two Approaches to Facilitate Gene Discovery. Genome Research 6, 791-806.]

TAG\_Lib=NCI\_CGAP\_Brn23  
TAG\_Tissue=brain  
TAG\_SEQ=ATATC

BASE COUNT 155 a 118 c 97 g 142 t  
ORIGIN

Query Match 100.0%; Score 263; DB 10; Length 512;  
Best Local Similarity 100.0%; Pred. No. 5e-61;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAACAATAACAGATTATTAGCTATATACAAAAAAGTCAGAGTAAGGCCAATC 60  
DB 18 AAACAATAACAGATTATTAGCTATATACAAAAAAGTCAGAGTAAGGCCAATC 77  
QY 61 TCAAGCAAGGCTTGATCTGCTACTTAACAATTTCACCAAGACTTCTCTTCCGCC 120  
DB 78 TCAAGCAAGGCTTGATCTGCTACTTAACAATTTCACCAAGACTTCTCTTCCGCC 137  
QY 121 TCTCAACTCTCCCTTCAGTGTGTCAGCTTACGATGTCCTGTCATGCCAAGGCC 180  
DB 138 TCTCAACTCTCCCTTCAGTGTGTCAGCTTACGATGTCCTGTCATGCCAAGGCC 197  
QY 181 CAAGGTGTCATCATTAAGACCCAGGAATCTACTACTCTTTTCATCAATCAAGGGGAA 240  
DB 198 CAAGGTGTCATCATTAAGACCCAGGAATCTACTACTCTTTTCATCAATCAAGGGGAA 257  
QY 241 TTTAAACAGCTTCTACCCAGCAT 263  
DB 258 TTTAAACAGCTTCTACCCAGCAT 280

RESULT 15  
BM676160  
LOCUS BM676160 522 bp mRNA linear EST 27-FEB-2002  
DEFINITION U1-E-EJ0-ahf-b-11-0-U1-s1 U1-E-EJ0 Homo sapiens cDNA clone  
ACCESSION U1-E-EJ0-ahf-b-11-0-U1-s1  
VERSION BM676160.1 GI:18986056  
KEYWORDS EST  
SOURCE human.

ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS 1 (bases 1 to 522)  
TITLE Normalization and subtraction: two approaches to facilitate gene discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
COMMENT Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: msoares@blue.weeg.uiowa.edu

FEATURES  
source  
1..522  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="U1-E-EJ0-ahf-b-11-0-U1"  
/clone\_lib="U1-E-EJ0"  
/tissue\_type="fetal eyes, lens, eye anterior segment, optic nerve, retina, Retina Paveal and Macular, RPE and Choroid"  
/dev\_stage="fetal and adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/note="Origin: eye; Vector: pRT73-Pac (Pharmacia) with a modified polylinker; Site\_1: EcoR I; Site\_2: Not I; U1-E-EJ0 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares. Genome Research, 6:797-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pRT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are: fetal eyes, ACATTCAGCA; lens, CGATTAGCGA; eye anterior segment, AATGCCGAT; optic nerve, CCATTAGTGTG; retina, CCGCG; Retina Paveal and Macular, GTCC; RPE and Choroid, ACCCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).  
TAG\_Lib=U1-E-EJ0  
TAG\_Tissue=human retina  
TAG\_SEQ=CCGCG"

BASE COUNT 154 a 124 c 99 g 144 t 1 others  
ORIGIN

Query Match 100.0%; Score 263; DB 14; Length 522;  
Best Local Similarity 100.0%; Pred. No. 5.1e-61;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAACAATAACAGATTATTAGCTATATACAAAAAAGTCAGAGTAAGGCCAATC 60  
DB 19 AAACAATAACAGATTATTAGCTATATACAAAAAAGTCAGAGTAAGGCCAATC 78  
QY 61 TCAAGCAAGGCTTGATCTGCTACTTAACAATTTCACCAAGACTTCTCTTCCGCC 120  
DB 79 TCAAGCAAGGCTTGATCTGCTACTTAACAATTTCACCAAGACTTCTCTTCCGCC 138  
QY 121 TCTCAACTCTCCCTTCAGTGTGTCAGCTTACGATGTCCTGTCATGCCAAGGCC 180

D <sub>b</sub>	139	TCTCAGCNCCTCCCTTCAGTGGGTGTCAGCTTCACGTTCTTCCTGGTGATGATGCCAAGGCC	198
O <sub>Y</sub>	181	CAGGTGTCATCATTAAGACCAGGAATCTACTACCTTTTTCACATTCAACAGGGAA	240
D <sub>b</sub>	199	CAGGTGTCATCATTAAGACCAGGAATCTACTACCTTTTTCACATTCAACAGGGAA	258
O <sub>Y</sub>	241	TTAAACACACTCTACTCCAGCAT	263
D <sub>b</sub>	259	TTAAACACACTCTACTCCAGCAT	281

Search completed: June 21, 2003, 03:37:16  
Job time : 788.299 secs



## ALIGNMENTS

for SEQ ID NO: 164

RESULT 1  
ABL67130  
ID ABL67130 standard; DNA: 451 BP.  
XX  
AC ABL67130;  
XX  
DT 15-MAY-2002 (first entry)  
XX  
DE Thyroid cancer related gene sequence SEQ ID NO:5467.  
XX  
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200194629-A2.  
XX  
PD 13-DEC-2001.  
XX  
PF 30-MAY-2001; 2001WO-US10838.  
XX  
PR 05-JUN-2000; 2000US-209473P.  
PR 05-JUN-2000; 2000US-209531P.  
PR 18-SEP-2000; 2000US-233133P.  
PR 18-SEP-2000; 2000US-233617P.  
PR 20-SEP-2000; 2000US-234009P.  
PR 20-SEP-2000; 2000US-234034P.  
PR 20-SEP-2000; 2000US-234052P.  
PR 22-SEP-2000; 2000US-234509P.  
PR 22-SEP-2000; 2000US-234567P.



for SET ID NO: 164

RESULT 1  
AA004887  
LOCUS  
DEFINITION 451 bp mRNA linear EST 23-JUL-1996  
2h90g01.s1 Soares fetal\_liver\_spleen\_INFLS\_S1 Homo sapiens cDNA

ACCESSION  
AA004887  
VERSION  
AA004887.1 GI:1447704  
KEYWORDS  
EST  
SOURCE  
human  
ORGANISM  
Homo sapiens

REFERENCE  
AUTHORS  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
1 (bases 1 to 451)  
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Tivanskis, E., Waterston, R., Williamson, A., Wohldmann, P. and Wilson, R.

TITLE  
JOURNAL  
COMMENT  
The WashU-Merck EST Project  
Unpublished (1995)  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.  
Seq primer: mob.REGA+ET  
High quality sequence stop: 429.

Mon Jun 23 10:00:35 2003

us-09-964-8:

FEATURES  
source Location/Qualifiers  
1..451  
/organism="Homo sapiens"  
/db\_xref="GDB:1328361"  
/db\_xref="taxon:9606"  
/clone="IMAGE:428592"  
/clone\_lib="Soares\_fetal\_liver\_spleen\_INFLS\_S1"  
/sex="male"  
/dev\_stage="20 week-post conception fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/note="Organ: Liver and Spleen; Vector: pTZ19 (Pharmacia) with a modified polylinker; Site 1: Pac I; Site 2: Eco RI; This is a subtracted version of the original Soares fetal liver spleen INFLS library. 1st strand cDNA was primed with a Pac I - 0190(dT) primer [5', AACTGAGAGATTAAAGATCTTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pTZ19 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 134 a 77 c 97 g 142 t 1 others

ORIGIN  
Query Match 99.8%; Score 450; DB 9; Length 451;  
Best Local Similarity 100.0%; Pred. No. 7,1e-108;  
Matches 451; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGGCCCTCCACTTTTATTTGAGTGTACATCTGTCCACTGTGCAAAATGAGATC 60  
DB 1 GCAGAGGCCCTCCACTTTTATTTGAGTGTACATCTGTCCACTGTGCAAAATGAGATC 60  
QY 61 ACAGCCTCACTCAATTCGAGAGGCCCTGCAAGNAAAGAAAGATGCCAGACAGTC 120  
DB 61 ACAGCCTCACTCAATTCGAGAGGCCCTGCAAGNAAAGAAAGATGCCAGACAGTC 120  
QY 121 TGTAGAGTTGATCTCGACTAATATCTTACAGCTTGTGAAATCACTGTCAAGGTT 180  
DB 121 TGTAGAGTTGATCTCGACTAATATCTTACAGCTTGTGAAATCACTGTCAAGGTT 180  
QY 181 TATTTAAATGAGATTTTGAAGATTAATTTTACGACTAATTTTAAATTAATCTAT 240  
DB 181 TATTTAAATGAGATTTTGAAGATTAATTTTACGACTAATTTTAAATTAATCTAT 240  
QY 241 GCAGAGTTGTTATTTGAGAGATTTGCCAAATTTAGAGCTGTGAGATGGAATTAATG 300  
DB 241 GCAGAGTTGTTATTTGAGAGATTTGCCAAATTTAGAGCTGTGAGATGGAATTAATG 300  
QY 301 CCTTCCTGACAGCTCTCTGTTTAAAGTGGTAAAGAAAGTTTCTTCCAGAAAAAT 360  
DB 301 CCTTCCTGACAGCTCTCTGTTTAAAGTGGTAAAGAAAGTTTCTTCCAGAAAAAT 360  
QY 361 ACAGCAGAAAAATCCGATGTTCTGATAGAGATTATTTGAGAGTGTGCCAGAGACAGCA 420  
DB 361 ACAGCAGAAAAATCCGATGTTCTGATAGAGATTATTTGAGAGTGTGCCAGAGACAGCA 420  
QY 421 GCTTCGTGATGTGACACCAATGTCTGT 451  
DB 421 GCTTCGTGATGTGACACCAATGTCTGT 451

## ALIGNMENTS

for seq ID NO: 174

## RESULT 1

ABL67140

ID ABL67140 standard; DNA; 272 BP.

XX

AC ABL67140;

XX

DT 15-MAY-2002 (first entry)

XX

DE Thyroid cancer related gene sequence SEQ ID NO:5477.

XX

KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.

XX

OS Homo sapiens.

XX

PN WO200194629-A2.

XX

PD 13-DEC-2001.

XX

PF 30-MAY-2001; 2001WO-US10838.

XX

PR 05-JUN-2000; 2000US-209473P.

PR 05-JUN-2000; 2000US-209531P.

PR 18-SEP-2000; 2000US-233133P.

PR 18-SEP-2000; 2000US-233617P.

PR 20-SEP-2000; 2000US-234009P.

PR 20-SEP-2000; 2000US-234034P.

PR 20-SEP-2000; 2000US-234052P.

PR 22-SEP-2000; 2000US-234509P.

PR 22-SEP-2000; 2000US-234567P.



6-552 10 ~0:174

Mon Jun 23 10:00:36 2003

us-09-964

RESULT 1  
AA281006  
LOCUS 272 bp mRNA linear EST 14-AUG-1997  
DEFINITION zs9406.s1 NCI\_CGAP\_GCB1 Homo sapiens cDNA clone IMAGE:705106 3',  
mRNA sequence.  
AA281006  
ACCESSION AA281006.1 GI:1923895  
VERSION  
KEYWORDS  
SOURCE EST.  
ORGANISM human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 272)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 845 Std Error: 0.00  
Seq primer: -41ml3 fwd. RT from Amersham  
High quality sequence stop: 252.  
FEATURES  
Source location/Qualifiers  
1..272  
/organism="Homo sapiens"  
/db\_xref="GDB:5854830"  
/db\_xref="taxon:9606"

BASE COUNT 45 a 57 c 47 g 123 t  
ORIGIN  
/clone="IMAGE:705106"  
/clone\_lib="NCI\_CGAP\_GCB1"  
/tissue\_type="germinal center B cell"  
/lab\_host="DH10B"  
/note="Vector: p773D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
was prepared from human tonsillar cells enriched for  
germinal center B cells by flow sorting (CD20+, IgD-),  
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman  
(NCI) and Dr. Gerald Marcil (CBER). cDNA synthesis was  
primed with a Not I - oligo(dT) primer  
[5'-TGTTACCAATCTGAAGTGGAGCGCGCCCTCATTTTCTTTT-3'  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified p773 vector. Library  
went through one round of normalization, and was  
constructed by Bento Soares and M. Fatima Bonaldo."

Query Match 100.0%; Score 272; DB 9; Length 272;  
Best Local Similarity 100.0%; Pred. No. 3e-45; Indels 0; Gaps 0;  
Matches 272; Conservative 0; Mismatches 0

QY	1	TTTTTTTTTAAATAGCTGCTCTCTGCATATAGTTTATTCCTTATCTTTTGAAC	60
DB	1	TTTTTTTTTAAATAGCTGCTCTCTGCATATAGTTTATTCCTTATCTTTTGAAC	60
QY	61	ATTTATACACCTTATTCATGTTCCCTTTAGATCACTATCTCTTACTCTCTGG	120
DB	61	ATTTATACACCTTATTCATGTTCCCTTTAGATCACTATCTCTTACTCTCTGG	120
QY	121	GCTTGAATCCTTGTCTGTTCTGATCTGCTGCTCTCTTGGATACCGGGAGTTT	180
DB	121	GCTTGAATCCTTGTCTGTTCTGATCTGCTGCTCTCTTGGATACCGGGAGTTT	180
QY	181	CCTCTGACACGCTCTTCAGTAGAGAAATGATTTTCATGAGATCTGTTCCCTGGATG	240
DB	181	CCTCTGACACGCTCTTCAGTAGAGAAATGATTTTCATGAGATCTGTTCCCTGGATG	240
QY	241	AGGACGCTCTCTCTGGGAGAAATGCTCTGT	272
DB	241	AGGACGCTCTCTCTGGGAGAAATGCTCTGT	272

## ALIGNMENTS

for SEQ ID NO: 180

RESULT 1  
ABL67146  
ID ABL67146 standard; DNA; 319 BP.  
XX  
AC ABL67146;  
XX  
DT 15-MAY-2002 (first entry)  
XX  
DE Thyroid cancer related gene sequence SEQ ID NO:5483.  
XX  
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200194629-A2.  
XX  
PD 13-DEC-2001.  
XX  
PF 30-MAY-2001; 2001WO-US10838.  
XX  
PR 05-JUN-2000; 2000US-209473P.  
PR 05-JUN-2000; 2000US-209531P.  
PR 18-SEP-2000; 2000US-233133P.  
PR 18-SEP-2000; 2000US-233617P.  
PR 20-SEP-2000; 2000US-234009P.  
PR 20-SEP-2000; 2000US-234034P.  
PR 20-SEP-2000; 2000US-234052P.  
PR 22-SEP-2000; 2000US-234509P.  
PR 22-SEP-2000; 2000US-234567P.

PR 25-SEP-2000; 2000US-234923P.  
PR 25-SEP-2000; 2000US-234924P.  
PR 25-SEP-2000; 2000US-235077P.  
PR 25-SEP-2000; 2000US-235082P.  
PR 25-SEP-2000; 2000US-235134P.  
PR 25-SEP-2000; 2000US-235280P.  
PR 26-SEP-2000; 2000US-235637P.  
PR 26-SEP-2000; 2000US-235638P.  
PR 27-SEP-2000; 2000US-235711P.  
PR 27-SEP-2000; 2000US-235720P.  
PR 27-SEP-2000; 2000US-235840P.  
PR 27-SEP-2000; 2000US-235863P.  
PR 28-SEP-2000; 2000US-236028P.  
PR 28-SEP-2000; 2000US-236032P.  
PR 28-SEP-2000; 2000US-236033P.  
PR 28-SEP-2000; 2000US-236034P.  
PR 28-SEP-2000; 2000US-236109P.  
PR 29-SEP-2000; 2000US-236111P.  
PR 29-SEP-2000; 2000US-236842P.  
PR 29-SEP-2000; 2000US-236891P.  
PR 02-OCT-2000; 2000US-237172P.  
PR 02-OCT-2000; 2000US-237173P.  
PR 02-OCT-2000; 2000US-237278P.  
PR 02-OCT-2000; 2000US-237278P.  
PR 02-OCT-2000; 2000US-237294P.  
PR 02-OCT-2000; 2000US-237295P.  
PR 03-OCT-2000; 2000US-237316P.  
PR 03-OCT-2000; 2000US-237425P.  
PR 03-OCT-2000; 2000US-237598P.  
PR 03-OCT-2000; 2000US-237604P.  
PR 03-OCT-2000; 2000US-237606P.  
PR 03-OCT-2000; 2000US-237608P.  
PR 01-NOV-2000; 2000US-244867P.  
PR 01-NOV-2000; 2000US-245084P.  
XX  
XX  
PA (AVALON PHARM.  
XX  
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
PI Soppet DR, Weaver Z;  
XX  
XX  
DR MPI: 2002-188264/24.  
XX  
XX  
PT Screening for anti-neoplastic agent involves exposing cells to a  
PT chemical agent to be tested for anti-neoplastic activity, and  
PT determining a change in expression of a gene of a signature gene set -  
XX  
PS Claim 1: SEQ ID 5483; 44pp; English.  
XX  
XX  
CC The present invention describes a method (M1) for screening for an  
CC anti-neoplastic agent. The method involves exposing cells to a chemical  
CC agent to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
CC to ABL70110), or is at least 95% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
CC activity and can be used in gene therapy. M1 can be used for screening  
CC an anti-neoplastic agent, and can be used for producing a product which  
CC is the data collected with respect to the anti-neoplastic agent as a  
CC result of M1, and the data is sufficient to convey the chemical  
CC structure and/or properties of the agent. M1 can be used in the  
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer.  
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
CC carcinoma, papillary carcinoma and Wilms' tumour.  
XX  
XX  
SQ Sequence 319 BP; 95 A; 73 C; 57 G; 94 T; 0 other;

Query Match 100.0%; Score 319; DB 24; Length 319;

Best Local Similarity 100.0%; Pred. No. 1.8e-85;

Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTTTCACGTGACCATGATTTAAATTTATTTAGTCCCAAAATGCTAGCTA 60  
|||||

DB 1 TTTTTCACGTGACCATGATTTAAATTTATTTAGTCCCAAAATGCTAGCTA 60  
QY 61 TTTTCAGTACATTTGATGAAACACCATTTCTTTATCTCTAAAGATGAGAAATTTGCT 120  
|||||  
DB 61 TTTTCAGTACATTTGATGAAACACCATTTCTTTATCTCTAAAGATGAGAAATTTGCT 120  
QY 121 ACTATATATTTTTCGCTATATCCAGCAGCAAGATACAAATGGAATCTCTATGATATTT 180  
|||||  
DB 121 ACTATATATTTTTCGCTATATCCAGCAGCAAGATACAAATGGAATCTCTATGATATTT 180  
QY 181 CTAAGCATATATGAGAAAGGCTCCAGGCTAAATGCAAGTATCTGATTAATGTTTC 240  
|||||  
DB 181 CTAAGCATATATGAGAAAGGCTCCAGGCTAAATGCAAGTATCTGATTAATGTTTC 240  
QY 241 CCCACACATGGAATCAACCTCCCGCTCCCTGGAAGCTTCCCAAGGTGCGGGGG 300  
|||||  
DB 241 CCCACACATGGAATCAACCTCCCGCTCCCTGGAAGCTTCCCAAGGTGCGGGGG 300  
QY 301 AAGCAGAGGAAAAAAGG 319  
|||||  
DB 301 AAGCAGAGGAAAAAAGG 319

## ALIGNMENTS

for seq ID no 180

RESULT 1  
AA490819  
LOCUS  
DEFINITION aa49f05.s1 NCI\_CGAP\_GCB1 319 bp mRNA linear EST 15-AUG-1997  
mRNA sequence.  
ACCESSION AA490819  
VERSION AA490819.1 GI:2219992  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 319)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,  
Ph.D., Gerald Marti, M.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
Bonaldo, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
Seq primer: -41m13 fwd. ET from Amersham.

FEATURES  
Source

Location/Qualifiers  
1. 319  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:824289"  
/clone\_lib="NCI\_CGAP\_GCB1"  
/tissue\_type="germinal center B cell"  
/lab\_host="DH10B"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr. David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGAGCGCGCTCATTTTTTTTTTTT-3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 95 a 73 c 57 g 94 t

Query Match 100.0%; Score 319; DB 9; Length 319;  
Best Local Similarity 100.0%; Pred. No. 1.9e-71;  
Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTTCTTCTCACTGTCCACCATGAATTTAAATTTATTTAGTGTCCACCAAAATGCTAGTCTA 60  
DB 1 TTTTCTTCTCACTGTCCACCATGAATTTAAATTTATTTAGTGTCCACCAAAATGCTAGTCTA 60

QY 61 TTCTCAGTACATTTGATGAACACCATTTCTTTATCTCTAAAGGATGAGAGATATTGCT 120  
DB 61 TTCTCAGTACATTTGATGAACACCATTTCTTTATCTCTAAAGGATGAGAGATATTGCT 120

QY 121 ACTATATATTTTTTTTGTCTATCATCCAGCCAGAGATACAAATGGAACCTCTATGAATATT 180  
DB 121 ACTATATATTTTTTTTGTCTATCATCCAGCCAGAGATACAAATGGAACCTCTATGAATATT 180

QY 181 CTAAGCATAATGAGGAGGCTCCAGGCTAAATGCAAGTATCTTAAAGGATGAGAGATATTGCT 240  
DB 181 CTAAGCATAATGAGGAGGCTCCAGGCTAAATGCAAGTATCTTAAAGGATGAGAGATATTGCT 240

QY 241 CCCACCTAGGGAATCACCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGG 300  
DB 241 CCCACCTAGGGAATCACCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGG 300

QY 301 AAGCAGGAGAGAAAAAAGG 319  
DB 301 AAGCAGGAGAGAAAAAAGG 319

## RESULT 2

LOCUS AA490870  
DEFINITION aa48b07.s1 NCI\_CGAP\_GCB1 Homo sapiens cDNA clone IMAGE:824149 3', mRNA linear EST 15-AUG-1997  
ACCESSION AA490870  
VERSION AA490870.1 GI:2220043  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 369)  
AUTHORS NCI-CGAP  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaps-ref@mail.nih.gov  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,

Ph.D., Gerald Marti, M.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
Clone Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/dbbrp/image/image.html  
Seq primer: -41m3 fwd. ET from Amersham  
High quality sequence stop: 327.

## FEATURES

## Source

Location/Qualifiers  
1. 369  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:824149"  
/clone\_lib="NCI\_CGAP\_GCB1"  
/tissue\_type="germinal center B cell"  
/lab\_host="DH10B"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr. David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGAGCGCGCTCATTTTTTTTTTTT-3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 116 a 83 c 67 g 103 t

Query Match 100.0%; Score 319; DB 9; Length 369;  
Best Local Similarity 100.0%; Pred. No. 1.9e-71;  
Matches 319; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTTCTTCTCACTGTCCACCATGAATTTAAATTTATTTAGTGTCCACCAAAATGCTAGTCTA 60  
DB 1 TTTTCTTCTCACTGTCCACCATGAATTTAAATTTATTTAGTGTCCACCAAAATGCTAGTCTA 60

QY 61 TTCTCAGTACATTTGATGAACACCATTTCTTTATCTCTAAAGGATGAGAGATATTGCT 120  
DB 61 TTCTCAGTACATTTGATGAACACCATTTCTTTATCTCTAAAGGATGAGAGATATTGCT 120

QY 121 ACTATATATTTTTTTTGTCTATCATCCAGCCAGAGATACAAATGGAACCTCTATGAATATT 180  
DB 121 ACTATATATTTTTTTTGTCTATCATCCAGCCAGAGATACAAATGGAACCTCTATGAATATT 180

QY 181 CTAAGCATAATGAGGAGGCTCCAGGCTAAATGCAAGTATCTTAAAGGATGAGAGATATTGCT 240  
DB 181 CTAAGCATAATGAGGAGGCTCCAGGCTAAATGCAAGTATCTTAAAGGATGAGAGATATTGCT 240

QY 241 CCCACCTAGGGAATCACCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGG 300  
DB 241 CCCACCTAGGGAATCACCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGG 300

QY 301 AAGCAGGAGAGAAAAAAGG 319  
DB 301 AAGCAGGAGAGAAAAAAGG 319

## ALIGNMENTS

for seq ID no: 191

RESULT 1  
ABL67157  
ID ABL67157 standard; DNA; 441 BP.  
XX  
AC ABL67157;  
XX  
DT 15-MAY-2002 (first entry)  
XX  
DE Thyroid cancer related gene sequence SEQ ID NO:5494.  
XX  
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200194629-A2.  
XX  
PD 13-DEC-2001.  
XX  
PF 30-MAY-2001; 2001WO-US10838.  
XX  
PR 05-JUN-2000; 2000US-209473P.  
PR 05-JUN-2000; 2000US-209531P.  
PR 18-SEP-2000; 2000US-233133P.  
PR 18-SEP-2000; 2000US-233617P.  
PR 20-SEP-2000; 2000US-234009P.  
PR 20-SEP-2000; 2000US-234034P.  
PR 20-SEP-2000; 2000US-234052P.  
PR 22-SEP-2000; 2000US-234509P.  
PR 22-SEP-2000; 2000US-234567P.

PR 25-SEP-2000; 2000US-234923P.  
 PR 25-SEP-2000; 2000US-234924P.  
 PR 25-SEP-2000; 2000US-235077P.  
 PR 25-SEP-2000; 2000US-235082P.  
 PR 25-SEP-2000; 2000US-235134P.  
 PR 25-SEP-2000; 2000US-235280P.  
 PR 26-SEP-2000; 2000US-235637P.  
 PR 26-SEP-2000; 2000US-235638P.  
 PR 27-SEP-2000; 2000US-235711P.  
 PR 27-SEP-2000; 2000US-235720P.  
 PR 27-SEP-2000; 2000US-235840P.  
 PR 27-SEP-2000; 2000US-235863P.  
 PR 28-SEP-2000; 2000US-236028P.  
 PR 28-SEP-2000; 2000US-236032P.  
 PR 28-SEP-2000; 2000US-236033P.  
 PR 28-SEP-2000; 2000US-236034P.  
 PR 28-SEP-2000; 2000US-236109P.  
 PR 28-SEP-2000; 2000US-236111P.  
 PR 29-SEP-2000; 2000US-236842P.  
 PR 29-SEP-2000; 2000US-236891P.  
 PR 02-OCT-2000; 2000US-237172P.  
 PR 02-OCT-2000; 2000US-237173P.  
 PR 02-OCT-2000; 2000US-237278P.  
 PR 02-OCT-2000; 2000US-237278P.  
 PR 02-OCT-2000; 2000US-237294P.  
 PR 02-OCT-2000; 2000US-237295P.  
 PR 02-OCT-2000; 2000US-237316P.  
 PR 03-OCT-2000; 2000US-237425P.  
 PR 03-OCT-2000; 2000US-237598P.  
 PR 03-OCT-2000; 2000US-237604P.  
 PR 03-OCT-2000; 2000US-237606P.  
 PR 03-OCT-2000; 2000US-237606P.  
 PR 01-NOV-2000; 2000US-244867P.  
 PR 01-NOV-2000; 2000US-245084P.  
 XX  
 PA (AVAL-) AVALON PHARM.  
 XX  
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrihan S;  
 PI Soppet DR, Weaver Z;  
 XX  
 DR WPI: 2002-188264/24.  
 XX  
 PT Screening for anti-neoplastic agent involves exposing cells to a  
 PT chemical agent to be tested for anti-neoplastic activity, and  
 PT determining a change in expression of a gene of a signature gene set -  
 XX  
 PS Claim 1: SEQ ID 5494; 44pp; English.  
 XX  
 CC The present invention describes a method (M1) for screening for an  
 CC anti-neoplastic agent. The method involves exposing cells to a chemical  
 CC agent to be tested for anti-neoplastic activity, determining a change in  
 CC expression of at least one gene (I) of a signature gene set, where (I)  
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL1664  
 CC to ABL70110), or is at least 95% identical to (S), where a change in  
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
 CC activity and can be used in gene therapy. M1 can be used for screening  
 CC an anti-neoplastic agent, and can be used for producing a product which  
 CC is the data collected with respect to the anti-neoplastic agent as a  
 CC result of M1, and the data is sufficient to convey the chemical  
 CC structure and/or properties of the agent. M1 can be used in the  
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
 CC carcinoma, papillary carcinoma and Wilm's tumour.  
 CC  
 XX  
 SQ Sequence 441 BP; 158 A; 68 C; 72 G; 143 T; 0 other;

Query Match 100.0%; Score 441; DB 24; Length 441;  
 Best Local Similarity 100.0%; Pred. No. 2,4e-85;  
 Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CATATATTAAAGCTGAATTTTATTACTAAATTAATCTATGCAAAAAAATTCGTGCC 60  
 |||

Db	1	CATATATTAAAGCTGAATTTTATTACTAAATTAATCTATGCAAAAAAATTCGTGCC	60
Qy	61	TGGCGTGAATTTTCACTCCATCAAGTGTATACATGATTTTTCATTTCATTACAAGCAG	120
Db	61	TGGCGTGAATTTTCACTCCATCAAGTGTATACATGATTTTTCATTTCATTACAAGCAG	120
Qy	121	GAGATGAATGTAGACAAAGTGTAGGAACATGGCAATTAATTAATTAATTTACAA	180
Db	121	GAGATGAATGTAGACAAAGTGTAGGAACATGGCAATTAATTAATTAATTTACAA	180
Qy	181	AAGCAAAAAAATTAACAGTGTACACATTTACTGAGATATATAAATAAGCAACAAC	240
Db	181	AAGCAAAAAAATTAACAGTGTACACATTTACTGAGATATATAAATAAGCAACAAC	240
Qy	241	AATCAATATATACAAAGTATTCGTTCTGTGTACTGAGATATCTATGTGACATC	300
Db	241	AATCAATATATACAAAGTATTCGTTCTGTGTACTGAGATATCTATGTGACATC	300
Qy	301	ATTCAACAAAAAAGTTCCTAATGAATGACATTTTGGAAATCATATGATCTCAGG	360
Db	301	ATTCAACAAAAAAGTTCCTAATGAATGACATTTTGGAAATCATATGATCTCAGG	360
Qy	361	GTTTAATCATTTAGGCTATTTACCGTTCCCTTTTGTAGTGAATTTATCCCAAGTGA	420
Db	361	GTTTAATCATTTAGGCTATTTACCGTTCCCTTTTGTAGTGAATTTATCCCAAGTGA	420
Qy	421	GATACGTCTCCAGGTGTAG 441	
Db	421	GATACGTCTCCAGGTGTAG 441	

6- 572 15 20:19 1

RESULT 1  
N71063 441 bp mRNA linear EST 14-MAR-1996  
LOCUS za86a11.s1 Soares\_fetal\_lung\_NbHL19w Homo sapiens cDNA clone  
DEFINITION IMAGE:299420 3', mRNA sequence.  
N71063  
VERSION N71063.1 GI:1227643  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE 1 (bases 1 to 441)  
AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, B., Rohlfing, T., Soares, M., Tan, F., Trevisan, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.  
TITLE The Mashu-Merc EST Project  
JOURNAL Unpublished (1995)  
COMMENT Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
Seq primer: m3 -40 forward  
High quality sequence stop: 344.

Mon Jun 23 10:00:37 2003

us-09-96

FEATURES		Location/Qualifiers	
source		1..441	
		/organism="Homo sapiens"	
		/db_xref="GDB:1244344"	
		/db_xref="taxon:9606"	
		/clone="IMAGE:299420"	
		/dev_stage="19 weeks"	
		/lab_host="DH10B (ampicillin resistant)"	
		/note="Organ: lung; Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGACGAGGAGCGCGCAATTTTTTTTTTTT-3'] double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and Eco RI sites of the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M. Fatima Bonaldo. This library was constructed from the same fetus as the fetal heart library. Soares fetal heart NBHL19w."	
BASE COUNT		158 a 68 c 72 g 143 t	
ORIGIN			
Query Match		100.0%; Score 441; DB 14; Length 441;	
Best Local Similarity		100.0%; Pred. No. 2.2e-75;	
Matches 441; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	CATTATTATAGCTGAATTTTATTTTACTTAATTAATGTCACAAAAAATCTGCGC	60
DB	1	CATTATTATAGCTGAATTTTATTTTACTTAATTAATGTCACAAAAAATCTGCGC	60
QY	61	TGCGGTGGAATTCACGCCATCAAGGTGACATGATTTTCTATTCTTCAACAACAG	120
DB	61	TGCGGTGGAATTCACGCCATCAAGGTGACATGATTTTCTATTCTTCAACAACAG	120
QY	121	GAGATGAATGTAGGACAGGTGTAGGAAACATGCAATTAATTAATTAATTAATTA	180
DB	121	GAGATGAATGTAGGACAGGTGTAGGAAACATGCAATTAATTAATTAATTAATTA	180
QY	181	AAGCAAAAAAATTAACAGGTGACACATTTATCTGATTAATTAATTAATTAATTA	240
DB	181	AAGCAAAAAAATTAACAGGTGACACATTTATCTGATTAATTAATTAATTAATTA	240
QY	241	AATCAATTAATTAACAGGTGATTTGTTCTGTGCTAGGATACCTATGACATTC	300
DB	241	AATCAATTAATTAACAGGTGATTTGTTCTGTGCTAGGATACCTATGACATTC	300
QY	301	ATTCAAAACAAAAAGTTCTTAATGAATGACATTTTGGGAAATCATATGATCTACG	360
DB	301	ATTCAAAACAAAAAGTTCTTAATGAATGACATTTTGGGAAATCATATGATCTACG	360
QY	361	GGTTAATCATTAAGGTGATTTACCGTTCCCTTTTATAGACTTATCCAGTGCA	420
DB	361	GGTTAATCATTAAGGTGATTTACCGTTCCCTTTTATAGACTTATCCAGTGCA	420
QY	421	GATAGTGTCCAGGTGAAG 441	
DB	421	GATAGTGTCCAGGTGAAG 441	

## ALIGNMENTS

RESULT 1  
ABL67165  
ID ABL67165 standard; DNA; 448 BP.  
XX  
AC ABL67165;  
XX  
DT 15-MAY-2002 (first entry)  
XX  
DE Thyroid cancer related gene sequence SEQ ID NO:5502.  
XX  
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200194629-A2.  
XX  
PD 13-DEC-2001.  
XX  
PF 30-MAY-2001; 2001WO-US10838.  
XX  
PR 05-JUN-2000; 2000US-209473P.  
PR 05-JUN-2000; 2000US-209531P.  
PR 18-SEP-2000; 2000US-233133P.  
PR 18-SEP-2000; 2000US-233617P.  
PR 20-SEP-2000; 2000US-234009P.  
PR 20-SEP-2000; 2000US-234034P.  
PR 20-SEP-2000; 2000US-234052P.  
PR 22-SEP-2000; 2000US-234509P.  
PR 22-SEP-2000; 2000US-234567P.

6- SEQ ID NO:199

PR	25-SEP-2000	2000US-23492.93
PR	25-SEP-2000	2000US-23493.34
PR	25-SEP-2000	2000US-23500.77
PR	25-SEP-2000	2000US-23507.77
PR	25-SEP-2000	2000US-23508.22
PR	25-SEP-2000	2000US-23513.14
PR	25-SEP-2000	2000US-23516.80
PR	26-SEP-2000	2000US-23563.77
PR	26-SEP-2000	2000US-23565.88
PR	27-SEP-2000	2000US-23571.11
PR	27-SEP-2000	2000US-23575.70
PR	27-SEP-2000	2000US-23584.04
PR	27-SEP-2000	2000US-23586.33
PR	28-SEP-2000	2000US-23602.82
PR	28-SEP-2000	2000US-23603.92
PR	28-SEP-2000	2000US-23603.93
PR	28-SEP-2000	2000US-23610.34
PR	28-SEP-2000	2000US-23610.34
PR	28-SEP-2000	2000US-23610.39
PR	28-SEP-2000	2000US-23611.11
PR	29-SEP-2000	2000US-23624.42
PR	29-SEP-2000	2000US-23689.11
PR	02-OCT-2000	2000US-23711.21
PR	02-OCT-2000	2000US-23711.21
PR	02-OCT-2000	2000US-23717.13
PR	02-OCT-2000	2000US-23727.84
PR	02-OCT-2000	2000US-23727.84
PR	02-OCT-2000	2000US-23729.52
PR	02-OCT-2000	2000US-23731.61
PR	03-OCT-2000	2000US-23745.52
PR	03-OCT-2000	2000US-23755.88
PR	03-OCT-2000	2000US-23760.44
PR	03-OCT-2000	2000US-23766.66
PR	03-OCT-2000	2000US-23766.66
PR	01-NOV-2000	2000US-24508.74
PR	01-NOV-2000	2000US-24508.74

(AVAL - ) AVALON PHARM

PI	Young PE,	Augustus M,	Carter KC,	Ebner R,	Endress G,	Horriqan S,
PI	Soppet DR,	Weaver Z,				

WPI; 2002-188264/24.

Claim 1; SEQ ID 5502; 44pp; English.

Claim 1; SEQ ID 5502; 44pp; English.

The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in AB161864 to AB170110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms' tumour.

50 Sequence 448 BP; 127 A; 89 C; 82 G; 150 T; 0 other;

Query Match	100.08;	Score 448;	DB 24;	Length 448;
-------------	---------	------------	--------	-------------

Matches	448;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
---------	------	--------------	----	------------	----	--------	----	------	----

Db	1	TTTTTTTTTTTTTATGTGACACCAATTTCTTATTTCATTTTTGGAGTTTCTGAAACG	60
Qy	61	AAAAATACAAATGATTTTCTGTATATGATCCTAGCCCTGACCTTGCTGAACTGATTAA	120
Db	61	AAAAATACAAATGATTTTCTGTATATGATCCTAGCCCTGACCTTGCTGAACTGATTAA	120
Qy	121	TTCTATTACATGATGATTTTTGTGTGGTTAGACCCCTTACACATCAAAATGAGTTTAA	180
Db	121	TTCTATTACATGATGATTTTTGTGTGGTTAGACCCCTTACACATCAAAATGAGTTTAA	180
Qy	181	AAAAAATGTCAGAGTGGGCCAGACCAACAACAGATGACAGTACGCTTTGCCCATACA	240
Db	181	AAAAAATGTCAGAGTGGGCCAGACCAACAACAGATGACAGTACGCTTTGCCCATACA	240
Qy	241	GAGATAAATTTAGTTTTTGGCAGTCCCTTCCCATAGAGATTGTATGSCAGTAGCAATTC	300
Db	241	GAGATAAATTTAGTTTTTGGCAGTCCCTTCCCATAGAGATTGTATGSCAGTAGCAATTC	300
Qy	301	ATGGCTACTGCCATACAACTGAACTGAAGTCCAGAAAGTTTATAGTGATCGGGCCACAG	360
Db	301	ATGGCTACTGCCATACAACTGAACTGAAGTCCAGAAAGTTTATAGTGATCGGGCCACAG	360
Qy	361	AGCTAATTTACTGTGGAGCCAAAGAAAGAAATTTATTCCTACTCTCTTGCCCACTAAGC	420
Db	361	AGCTAATTTACTGTGGAGCCAAAGAAAGAAATTTATTCCTACTCTCTTGCCCACTAAGC	420
Qy	421	TCCCAATTCAGTGGGCTGCTTTCTGGT	448
Db	421	TCCCAATTCAGTGGGCTGCTTTCTGGT	448

OY	421	TCCCCATTCCAGTGGGCTGCTTCTGTG	448
Db	421	TCCCCATTCCAGTGGGCTGCTTCTGTG	448

for 522 10 Nov 1995

**US-09-964.**

FEATURES	Location/Qualifiers
source	1. .448
	/organism="Homo sapiens"
	/abstract=non-free

```

/organism="Homo sapiens"
/db_xref="GDB:556253"
/db_xref="taxon:9606"
/clone="IMAGE:66856"
/clone_1p=" Soares NHNPu.51"
/tissue_type="Pooled human melanocyte fetal heart cell"

```

```
/tissue_type="Pooled human melanocyte, fetal heart, and  
pregnant uterus"  
/lab_host="DHIOB"  
/note="Organ: mixed (see below); Vector: pT73D-Pac  
(Pharmacia) with a modified polylinker; Site: 1; Not I;  
Site 2: Eco RI; Equal amounts of plasmid DNA from three  
normalized libraries (melanocyte 2NMW, pregnant uterus
```

BASE COUNT	127 a	89 c	82 g	150 t
ORIGIN				

Query Match	Similarity	Score	DB %	Length
Best Local	Similarity	100.0%	DB 9%	Length 448:
Matches	448:	Conservative	0:	Mismatches 73; Indels 0; Gaps 0;
QY	1	TTTTTTTTTTTATTTGACACAAATTTCTTAATTCATTTTGGATTTTGTGACAG	60	
Db	1	TTTTTTTTTTTATTTGACACAAATTTCTTAATTCATTTTGGATTTTGTGACAG	60	
QY	61	AAAAATCAATTTGATTTTCTGTATTTGATGTATACCTGTGACTGTAACCTGATTA	120	
Db	61	AAAAATCAATTTGATTTTCTGTATTTGATGTATACCTGTGACTGTAACCTGATTA	120	
QY	121	TTCTATTACACTAAGATTTTGTGTGGTGTAGACCTTACACATCAATGAGGTTTAA	180	
Db	121	TTCTATTACACTAAGATTTTGTGTGGTGTAGACCTTACACATCAATGAGGTTTAA	180	
QY	181	AAAAATGTGCAGAGTGGCCCCAGACCAACACAGATGACATGACCTTTGCCATACA	240	
Db	181	AAAAATGTGCAGAGTGGCCCCAGACCAACAGATGACATGACCTTTGCCATACA	240	
QY	241	GAGATAAATTAATGATTTTGGAGTCTTTCCCATAGAGATGTATGGCAGTAGAATCT	300	
Db	241	GAGATAAATTAATGATTTTGGAGTCTTTCCCATAGAGATGTATGGCAGTAGAATCT	300	
QY	301	ATGGCTTACTGCCATACACACCTGAACTGAAAGTTCAGAAAGTTTATGAGTACGGGCCACAG	360	
Db	301	ATGGCTTACTGCCATACACACCTGAACTGAAAGTTCAGAAAGTTTATGAGTACGGGCCACAG	360	
QY	361	AGCTAATTAATGAGTGGAGCCAGAGAGAAATTTATTCCTTACCTCTTGCCCACTAAGC	420	
Db	361	AGCTAATTAATGAGTGGAGCCAGAGAGAAATTTATTCCTTACCTCTTGCCCACTAAGC	420	
QY	421	TTCCCATTCACAGTGGGCTCTTCTGCT	448	
Db	421	TTCCCATTCACAGTGGGCTCTTCTGCT	448	

## ALIGNMENTS

RESULT 1  
ABL67310  
ID ABL67310 standard; DNA; 283 BP.  
XX  
AC ABL67310;  
XX  
DT 15-MAY-2002 (first entry)  
XX  
DE Thyroid cancer related gene sequence SEQ ID NO:5647.  
XX  
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200194629-A2.  
XX  
PD 13-DEC-2001.  
XX  
PF 30-MAY-2001; 2001WO-US10838.  
XX  
PR 05-JUN-2000; 2000US-209473P.  
PR 05-JUN-2000; 2000US-209531P.  
PR 18-SEP-2000; 2000US-233133P.  
PR 18-SEP-2000; 2000US-233617P.  
PR 20-SEP-2000; 2000US-234009P.  
PR 20-SEP-2000; 2000US-234034P.  
PR 20-SEP-2000; 2000US-234052P.  
PR 22-SEP-2000; 2000US-234509P.  
PR 22-SEP-2000; 2000US-234567P.

For SEQ ID NO:344



6- SEP 15 NO 344

Mon Jun 23 10:00:39 2003

us-09-96

RESULT 1  
AA411711 283 bp mRNA linear EST 17-MAY-1997  
LOCUS zvl6d08.s1 Soares\_NHMPu\_S1 Homo sapiens cDNA clone IMAGE:753807  
DEFINITION 3', mRNA sequence.  
ACCESSION AA411711  
VERSION AA411711.1 GI:2069500  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 283)  
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S.,  
Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,  
'T., Waterston, R., and Wilson, R.  
WashU-Merck EST Project 1997  
Unpublished (1997)  
Contact: Wilson R  
Washington University School of Medicine  
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Possible reversed clone: polyA not found  
Seq primer: -41ml3 fwd. ET from Amersham.

FEATURES	Location/Qualifiers
source	1..283 /organism="Homo sapiens" /db_xref="GDB:5976674" /db_xref="taxon:9606" /clone="IMAGE:753807" /tissue_type="Soares_NHMPu_S1" /pregnant_uterus="Pooled human melanocyte, fetal heart, and /lab_host="DH10B" /note="Organ: mixed (see below); Vector: pT7T3p-Pac (Pharmacia) with a modified polylinker; Site_1: Not 1; Site_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NDH, pregnant uterus NbHPV, and fetal heart 2NDH) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of 1 M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."
BASE COUNT	63 a 95 c 90 g 35 t
ORIGIN	
Query Match	100.0%; Score 283; DB 9; Length 283;
Best Local Similarity	100.0%; Pred. No. 6; 6e-63;
Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 GCAGCCGCTCTTAAGACCTGCTGCTGCTGCCGCAAGCCCAAGAGCAGCTGTGGT 60
DB	1 GCAGCCGCTCTCTAAGAACCTGCTGCTGCTGCCGCAAGCCCAAGAGCAGCTGTGGT 60
QY	61 GCGGCCGAGAGCCTGTGGCAGGGGACATGGCATGCCAGAGTAAAAAACGACGGCG 120
DB	61 GCGGCCGAGAGCCTGTGGCAGGGGACATGGCATGCCAGAGTAAAAAACGACGGCG 120
QY	121 CGGAACAGAACTGTCATCTCCCGAGCCATCCTATGCAAGCAGCCCAAGCAGCAGAG 180
DB	121 CGGAACAGAACTGTCATCTCCCGAGCCATCCTATGCAAGCAGCCCAAGCAGCAGAG 180
QY	181 CCGAGTACTCAGAGCTGCTGGCCAAAGCTGCTTCCCTGAACGCCAGAGCAGTGGCTGG 240
DB	181 CCGAGTACTCAGAGCTGCTGGCCAAAGCTGCTTCCCTGAACGCCAGAGCAGTGGCTGG 240
QY	241 ACGGTGCTCACCAGCCGCTGCTGGAGCACCAGAGGAGCGCGAG 283
DB	241 ACGGTGCTCACCAGCCGCTGCTGGAGCACCAGAGGAGCGCGAG 283

PT involves detecting the level of expression of two or more genes in a liver tissue sample

Claim 1: SEQ ID NO 2679; 298bp; English.

The invention relates to a novel method for diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma or metastatic liver tumour in a patient, and differentiating metastatic liver cancer from hepatocellular carcinoma in a patient, involving detecting the level of expression of two or more genes represented in ABN9203-ABN97455 in a tissue sample. The method of the invention has hepatotrophic, and cytostatic activity. The method is useful for diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma and metastatic liver carcinoma in a patient. The method is useful for identifying expression profiles which serve as useful diagnostic markers as well as markers that can be used to monitor disease states, disease progression, drug toxicity, drug efficacy and drug metabolism.

CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [http://wipo.int/pub/published\\_pat\\_sequences](http://wipo.int/pub/published_pat_sequences).

XX Sequence 165 BP; 68 A; 22 C; 23 G; 52 T; 0 other;

Query Match 100.0%; Score 165; DB 24; Length 165;  
Best Local Similarity 100.0%; Pred. No. 3,7e-27;  
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCCTTGAGTAACCTTATTTTGGAGAGTTCATTAAGCATTAGACATACATTAATGA 60  
DB 1 CTCCTTGAGTAACCTTATTTTGGAGAGTTCATTAAGCATTAGACATACATTAATGA 60  
DB 61 CACACCACTGTGACATGAAAAAAGACATTTGATTTTCCAGCTTTTAAGTT 120  
DB 61 CACACCACTGTGACATGAAAAAAGACATTTGATTTTCCAGCTTTTAAGTT 120  
QY 121 AAAAAATGATTCAGTTAAACAAAAACAAGTTTATGATATTTTAA 165  
DB 121 AAAAAATGATTCAGTTAAACAAAAACAAGTTTATGATATTTTAA 165

RESULT 2  
ABU62688  
ID ABU62688 standard; DNA; 165 BP.

XX ABL62688;

XX 15-MAY-2002 (first entry)

DE Colon adenocarcinoma related gene sequence SEQ ID NO:1025.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
XX stomach; lung; prostate; carcinoma; antitumour; cancerous;  
XX cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;  
XX gene; ds.

OS Homo sapiens.

PN WO200194629-A2.

PD 13-DEC-2001.

PF 30-MAY-2001; 2001WO-US10838.

PR 05-JUN-2000; 2000US-209473P.

PR 05-JUN-2000; 2000US-209531P.

PR 18-SEP-2000; 2000US-233133P.

PR 18-SEP-2000; 2000US-233617P.

PR 20-SEP-2000; 2000US-234009P.

PR 20-SEP-2000; 2000US-234034P.

PR 20-SEP-2000; 2000US-234052P.

PR 22-SEP-2000; 2000US-234509P.

PR 22-SEP-2000; 2000US-234567P.

PR 25-SEP-2000; 2000US-234923P.

PR 25-SEP-2000; 2000US-234924P.  
PR 25-SEP-2000; 2000US-235077P.  
PR 25-SEP-2000; 2000US-235082P.  
PR 25-SEP-2000; 2000US-235134P.  
PR 25-SEP-2000; 2000US-235280P.  
PR 25-SEP-2000; 2000US-235637P.  
PR 26-SEP-2000; 2000US-235638P.  
PR 26-SEP-2000; 2000US-235711P.  
PR 27-SEP-2000; 2000US-235720P.  
PR 27-SEP-2000; 2000US-235840P.  
PR 27-SEP-2000; 2000US-235863P.  
PR 28-SEP-2000; 2000US-236028P.  
PR 28-SEP-2000; 2000US-236032P.  
PR 28-SEP-2000; 2000US-236033P.  
PR 28-SEP-2000; 2000US-236034P.  
PR 28-SEP-2000; 2000US-236109P.  
PR 28-SEP-2000; 2000US-236111P.  
PR 29-SEP-2000; 2000US-236842P.  
PR 29-SEP-2000; 2000US-236891P.  
PR 29-SEP-2000; 2000US-237172P.  
PR 02-OCT-2000; 2000US-237173P.  
PR 02-OCT-2000; 2000US-237278P.  
PR 02-OCT-2000; 2000US-237294P.  
PR 02-OCT-2000; 2000US-237295P.  
PR 02-OCT-2000; 2000US-237316P.  
PR 03-OCT-2000; 2000US-237425P.  
PR 03-OCT-2000; 2000US-237598P.  
PR 03-OCT-2000; 2000US-237604P.  
PR 03-OCT-2000; 2000US-237606P.  
PR 03-OCT-2000; 2000US-237608P.  
PR 03-OCT-2000; 2000US-237608P.  
PR 01-NOV-2000; 2000US-244867P.  
PR 01-NOV-2000; 2000US-245084P.

(AVAL-) AVALON PHARM.

PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;

PI Soppet DR, Weaver Z;

DR WPI: 2002-188264/24.

PT Screening for anti-neoplastic agent involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, and determining a change in expression of a gene of a signature gene set

PS Claim 1: SEQ ID 1025; 44pp; English.

XX The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (1) of a signature gene set, where (1) comprises a sequence (S) selected from 8447 sequences (given in ABU61664 to ABU70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (1) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms' tumour.

XX Sequence 165 BP; 68 A; 22 C; 23 G; 52 T; 0 other;

Query Match 100.0%; Score 165; DB 24; Length 165;  
Best Local Similarity 100.0%; Pred. No. 3,7e-27;  
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCCTTGAGTAACCTTATTTTGGAGAGTTCATTAAGCATTAGACATACATTAATGA 60  
DB 1 CTCCTTGAGTAACCTTATTTTGGAGAGTTCATTAAGCATTAGACATACATTAATGA 60

```

OY 61 CACACACTGTTGACATGAAAAAAGACATTGATATTTCCAGCTTTTAACTT 120
DB 61 CACACCACTGTTGACAAATGAAAAAAGACATTGATATTTCCAGCTTTTAACTT 120
OY 121 AAAAAATGATTCAGTTAAACAAAACAAAGTTAGATATTTAG 165
DB 121 AAAAAATGATTCAGTTAAACAAAACAAAGTTAGATATTTTTC 165

RESULT 3
ABL67347
ID ABL67347 standard: DNA: 165 BP.
XX
AC ABL67347:
XX
DT 15-MAY-2002 (first entry)
XX
DE Thyroid cancer related gene sequence SEQ ID NO:5684.
XX
KW Human; Cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
XX gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001: 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235072P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237588P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.

```

```

PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX WPI: 2002-188264/24.
XX
PR Screening for anti-neoplastic agent involves exposing cells to a
PR chemical agent to be tested for anti-neoplastic activity, and
PR determining a change in expression of a gene of a signature gene set
XX
XX Claim 1; SEQ ID 5684; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an
XX anti-neoplastic agent. The method involves exposing cells to a chemical
XX agent to be tested for anti-neoplastic activity, determining a change in
XX expression of at least one gene (I) of a signature gene set, where (I)
XX comprises a sequence (S) selected from 8447 sequences (given in ABL61664
XX to ABL70110), or is at least 95% identical to (S), where a change in
XX expression is indicative of anti-neoplastic activity. (I) has cytostatic
XX activity and can be used in gene therapy. M1 can be used for screening
XX an anti-neoplastic agent, and can be used for producing a product which
XX is the data collected with respect to the anti-neoplastic agent as a
XX result of M1, and the data is sufficient to convey the chemical
XX structure and/or properties of the agent. M1 can be used in the
XX treatment of cancer such as colon, breast, stomach, lung, thyroid,
XX oesophageal, ovarian, kidney, prostate or pancreatic cancer,
XX adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
XX infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
XX carcinoma, papillary carcinoma and Wilm's tumour.
XX
XX Sequence 165 BP; 68 A; 22 C; 23 G; 52 T; 0 other;
XX
XX Query Match 100.0%; Score 165; DB 24; Length 165;
XX Best Local Similarity 100.0%; Pred. No. 3.7e-27;
XX Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 CTCTTGAGTAAGTATTTTGGAGAGCTTCATTAAGCATTAAGCAATACATAAATGA 60
DB 1 CTCTTGAGTAAGTATTTTGGAGAGCTTCATTAAGCATTAAGCAATACATAAATGA 60
OY 61 CACACCACTGTTGACAAATGAAAAAAGACATTGATATTTCCAGCTTTTAACTT 120
DB 61 CACACCACTGTTGACAAATGAAAAAAGACATTGATATTTCCAGCTTTTAACTT 120
OY 121 AAAAAATGATTCAGTTAAACAAAACAAAGTTAGATATTTAG 165
DB 121 AAAAAATGATTCAGTTAAACAAAACAAAGTTAGATATTTTTC 165

RESULT 4
ABL69408
ID ABL69408 standard: DNA: 165 BP.
XX
AC ABL69408:
XX
DT 15-MAY-2002 (first entry)
XX
DE Prostate cancer related gene sequence SEQ ID NO:7745.
XX
KW Human; Cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
XX gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.

```

XX 30-MAY-2001; 2001WO-US10838.  
 PF 05-JUN-2000; 2000US-209473P.  
 PR 05-JUN-2000; 2000US-209531P.  
 PR 18-SEP-2000; 2000US-233133P.  
 PR 18-SEP-2000; 2000US-233617P.  
 PR 20-SEP-2000; 2000US-234009P.  
 PR 20-SEP-2000; 2000US-234034P.  
 PR 20-SEP-2000; 2000US-234052P.  
 PR 22-SEP-2000; 2000US-234509P.  
 PR 22-SEP-2000; 2000US-234567P.  
 PR 25-SEP-2000; 2000US-234923P.  
 PR 25-SEP-2000; 2000US-234924P.  
 PR 25-SEP-2000; 2000US-235077P.  
 PR 25-SEP-2000; 2000US-235082P.  
 PR 25-SEP-2000; 2000US-235134P.  
 PR 25-SEP-2000; 2000US-235280P.  
 PR 26-SEP-2000; 2000US-235637P.  
 PR 26-SEP-2000; 2000US-235638P.  
 PR 27-SEP-2000; 2000US-235711P.  
 PR 27-SEP-2000; 2000US-235720P.  
 PR 27-SEP-2000; 2000US-235840P.  
 PR 27-SEP-2000; 2000US-235863P.  
 PR 28-SEP-2000; 2000US-236028P.  
 PR 28-SEP-2000; 2000US-236032P.  
 PR 28-SEP-2000; 2000US-236033P.  
 PR 28-SEP-2000; 2000US-236034P.  
 PR 28-SEP-2000; 2000US-236109P.  
 PR 28-SEP-2000; 2000US-236111P.  
 PR 29-SEP-2000; 2000US-236842P.  
 PR 29-SEP-2000; 2000US-236891P.  
 PR 02-OCT-2000; 2000US-237172P.  
 PR 02-OCT-2000; 2000US-237173P.  
 PR 02-OCT-2000; 2000US-237278P.  
 PR 02-OCT-2000; 2000US-237295P.  
 PR 02-OCT-2000; 2000US-237295P.  
 PR 02-OCT-2000; 2000US-237316P.  
 PR 03-OCT-2000; 2000US-237425P.  
 PR 03-OCT-2000; 2000US-237598P.  
 PR 03-OCT-2000; 2000US-237604P.  
 PR 03-OCT-2000; 2000US-237606P.  
 PR 03-OCT-2000; 2000US-237608P.  
 PR 01-NOV-2000; 2000US-244867P.  
 PR 01-NOV-2000; 2000US-245084P.  
 PA (AVAL-) AVALON PHARM.  
 XX  
 XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S,  
 PI Soppet DR, Weaver Z;  
 XX  
 XX WPI: 2002-188264/24.  
 DR  
 XX  
 XX Screening for anti-neoplastic agent involves exposing cells to a  
 PT chemical agent to be tested for anti-neoplastic activity, and  
 PT determining a change in expression of a gene of a signature gene set -  
 XX  
 PS Claim 1: SEQ ID 7745: 44pp; English.  
 XX  
 CC The present invention describes a method (M1) for screening for an  
 CC anti-neoplastic agent. The method involves exposing cells to a chemical  
 CC agent to be tested for anti-neoplastic activity, determining a change in  
 CC expression of at least one gene (I) of a signature gene set, where (I)  
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
 CC to ABL70110), or is at least 95% identical to (S), where a change in  
 CC expression is indicative of anti-neoplastic activity. (I) has cytoskeletal  
 CC activity and can be used in gene therapy. M1 can be used for screening  
 CC an anti-neoplastic agent, and can be used for producing a product which  
 CC is the data collected with respect to the anti-neoplastic agent as a  
 CC result of M1, and the data is sufficient to convey the chemical  
 CC structure and/or properties of the agent. M1 can be used in the  
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
 CC lymphoma, ovarian, kidney, prostate or pancreatic cancer,

CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
 CC carcinoma, papillary carcinoma and Wilms' tumour.  
 XX  
 SQ Sequence 165 BP; 68 A; 22 C; 23 G; 52 T; 0 other;  
 Query Match 100.0%; Score 165; DB 24; Length 165;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-27;  
 Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CTCCTTGAGTAACTTATTTTGGAGGAGTTCATTAAGCATTTAGCAACATCATTAATAATGA 60  
 1 CTCCTTGAGTAACTTATTTTGGAGGAGTTCATTAAGCATTTAGCAACATCATTAATAATGA 60  
 Db 61 CACACCACGTTTACATGAAAAAACAAGCATTTGATATTTTCAGCTTTTAAAGTT 120  
 61 CACACCACGTTTACATGAAAAAACAAGCATTTGATATTTTCAGCTTTTAAAGTT 120  
 Db 61 CACACCACGTTTACATGAAAAAACAAGCATTTGATATTTTCAGCTTTTAAAGTT 120  
 QY 121 AAAAAATGATTCAGTAAAAACAACAAGTTAGATATTTTAAAG 165  
 121 AAAAAATGATTCAGTAAAAACAACAAGTTAGATATTTTAAAG 165  
 Db 121 AAAAAATGATTCAGTAAAAACAACAAGTTAGATATTTTAAAG 165  
 RESULT 5  
 ID AAF91861/c  
 ID AAF91861 standard; cDNA; 5330 BP.  
 AC AAF91861;  
 XX  
 XX 22-MAY-2001 (first entry)  
 DT  
 XX  
 DE Human secreted protein-encoding gene 4 cDNA clone HDPFY41, SEQ ID NO:14.  
 XX  
 KW Human; secreted protein; proliferative disorder; cancer; tumour;  
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;  
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;  
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;  
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;  
 KW cardiovascular disorder; sepsis; diabetes; atherosclerosis;  
 KW gastrointestinal disorder; angioedema; kidney disorder;  
 KW endocrine disorder; infection; pregnancy-related disorder;  
 KW cell culture; chemotaxis; food additive;  
 KW binding partner identification; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200118022-A1.  
 XX  
 PD 15-MAR-2001.  
 XX  
 PF 31-AUG-2000; 2000WO-US24008.  
 XX  
 PR 03-SEP-1999; 99US-0152315.  
 PR 03-SEP-1999; 99US-0152317.  
 XX  
 PA (HUMA-) HUMA GENOME SCI INC.  
 XX  
 PI Ni J, Baker KP, Birse CE, Piscella M, Komatsoulis GA, Rosen CA;  
 PI Soppet DR, Young PE, Ebner R, Duan DR, Olsen HS, Lafleur DW;  
 PI Moore PA, Shi Y, Wei Y, Florence KA;  
 XX  
 XX WPI: 2001-203081/20.  
 DR P-PDB: AAB87345.  
 XX  
 PT Nucleic acid molecules encoding human secreted proteins, used in  
 PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and  
 PT Parkinson's diseases and cancers -  
 XX  
 PS Claim 1: Page 490-491; 607pp; English.  
 CC AAF91858-AAF91929 represent cDNAs corresponding to 52 human secreted  
 CC protein genes, and AAB87345-AAB87413 represent the proteins they encode.

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 20, 2003, 20:42:18 ; Search time 488.851 Seconds  
(without alignments)  
5466.401 Million cell updates/sec

Title: US-09-964-824C-381

Perfect score: 165  
Sequence: 1 cctcttgagtaacttattt.....caaaagttagatatlttag 165

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST: \*  
1: em\_estba: \*  
2: em\_esthum: \*  
3: em\_estlin: \*  
4: em\_estmu: \*  
5: em\_estov: \*  
6: em\_estopl: \*  
7: em\_estro: \*  
8: em\_hnc: \*  
9: gb\_est1: \*  
10: gb\_est2: \*  
11: gb\_hnc: \*  
12: gb\_est3: \*  
13: gb\_est4: \*  
14: gb\_est5: \*  
15: em\_estfun: \*  
16: em\_estom: \*  
17: gb\_gss: \*  
18: em\_gss\_hum: \*  
19: em\_gss\_inv: \*  
20: em\_gss\_pln: \*  
21: em\_gss\_vrt: \*  
22: em\_gss\_fun: \*  
23: em\_gss\_man: \*  
24: em\_gss\_mus: \*  
25: em\_gss\_other: \*  
26: em\_gss\_pro: \*  
27: em\_gss\_rtd: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	165	100.0	165	14	N73808	N73808.yz80g07.s1
2	165	100.0	213	9	A1699181	A1699181.tx61e1.x
3	165	100.0	282	14	R50866	R50866.yg62h08.s1
4	165	100.0	302	9	A1910763	A1910763.wg96h11.x
5	165	100.0	303	9	A1969490	A1969490.wz67f01.x
6	165	100.0	336	9	A1351615	A1351615.gr07f12.x

7	165	100.0	381	9	AA043474	AA043474.zk54e09.s
8	165	100.0	412	14	H08164	H08164.y187e01.s1
9	165	100.0	423	9	A1049699	A1049699.an33d09.x
10	165	100.0	423	10	AM592865	AM592865.hg04b07.x
11	165	100.0	444	9	A1041596	A1041596.ox56h06.x
12	165	100.0	450	10	AM072470	AM072470.xa60d07.x
13	165	100.0	451	9	A1573107	A1573107.tr69f01.x
14	165	100.0	459	10	AM770384	AM770384.h180d09.x
15	165	100.0	462	9	AA479302	AA479302.zv21f08.s
16	165	100.0	484	10	AM512118	AM512118.xx70f03.x
17	165	100.0	540	9	A1768486	A1768486.wh2e02.x
18	165	100.0	545	9	A1857643	A1857643.wk95h01.x
19	165	100.0	549	12	BF061430	BF061430.7j53f07.x
20	163	98.8	502	14	N48185	N48185.yz03d10.s1
21	154	93.3	234	10	AM150540	AM150540.x052h03.x
22	153	93.2	368	9	AA976261	AA976261.oq67a06.s
23	153	92.7	408	9	A1873695	A1873695.wm2e08.x
24	153	92.7	536	12	BG114119	BG114119.602285531
25	152	92.1	390	9	AA281459	AA281459.zs96d03.s
26	152	92.1	459	10	AM444556	AM444556.ui-h-B13-
27	151.4	91.8	387	9	A1915793	A1915793.wg94a07.x
28	151.4	91.8	568	12	BE879247	BE879247.601491650
29	151	91.5	542	9	A1679751	A1679751.tu76c08.x
30	150.4	91.2	505	9	A1693720	A1693720.wd44e08.x
31	149.2	90.4	511	9	AA527097	AA527097.n126e02.s
32	148.8	90.2	188	9	AA743364	AA743364.ny18d01.s
33	143	86.7	169	10	AM613593	AM613593.hb79f09.x
34	140	84.8	393	10	AM183811	AM183811.xj88b10.x
35	140	84.8	698	14	BM997974	BM997974.ui-h-DT1-
36	136.8	82.9	690	12	BG289398	BG289398.602381427
37	136	82.4	358	9	AA909938	AA909938.om12h12.s
38	134.6	81.6	561	10	AM293978	AM293978.ui-h-BM0-
39	112	67.9	888	9	AL561910	AL561910.AL561910
40	104.2	63.2	256	9	AV152678	AV152678.AV152678
41	104.2	63.2	301	10	BB070199	BB070199.BB070199
42	104.2	63.2	363	10	BE335251	BE335251.ug95h07.x
43	104.2	63.2	392	12	BF450263	BF450263.uz67e01.x
44	104.2	63.2	425	14	BM899961	BM899961.ui-m-DJ1-
45	104.2	63.2	427	17	AZ630834	AZ630834.lm0484B12

#### ALIGNMENTS

RESULT 1  
LOCUS N73808 165 bp mRNA linear EST 19-MAR-1996  
DEFINITION yz80g07.s1 Soares,multiple\_sclerosis\_2NDHMS Homo sapiens cDNA  
Clone IMAGE:289404 3', mRNA sequence.  
ACCESSION N73808  
VERSION N73808.1 GI:1231093  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 165)  
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Ritzin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Woldmann,P. and Wilson,R.  
The Washu-Merck EST Project  
Unpublished (1995)  
TITLE JOURNAL  
COMMENT Contact: Wilson RK  
Washington University School of Medicine  
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.wustl.edu  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (infoimage.llnl.gov) for further information.  
Seq primer: ml3 -40 forward  
High quality sequence stop: 144.



```
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/notes="Organ: whole brain; Vector: latmid BA; Site: 1: Not
1: Site 2: Hind III; 1st strand cDNA was primed with a Not
1 - oligo(dT) primer 15',
AAGTGGAGAAATTCGGCCGCGCAGCAATTTTCTTTTCTTTT 3';
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the latmid BA vector.
Library went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo."
```

BASE COUNT 101 a 43 c 40 g 96 t 2 others

ORIGIN

Query Match 100.0%; Score 165; DB 14; Length 282;  
Best Local Similarity 100.0%; Pred. No. 1.3e-21;  
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCTTGAGTAAGTATTTTGGAGAGTTCATAGCAATGGAACATACATAAATGA 60  
|||||  
DB 10 CTCTTGAGTAAGTATTTTGGAGAGTTCATAGCAATGGAACATACATAAATGA 69  
QY 61 CACACCACTGTTGACAAATGAAAAAAGACATTTGATATTTTCACCTTTTAACTT 120  
|||||  
DB 70 CACACCACTGTTGACAAATGAAAAAAGACATTTGATATTTTCACCTTTTAACTT 129  
QY 121 AAAAAATGATTCAGTTAAACAAACAAAGTTGATTTTAACTT 165  
|||||  
DB 130 AAAAAATGATTCAGTTAAACAAACAAAGTTGATTTTAACTT 174

RESULT 4  
AI910763 302 bp mRNA linear EST 20-DEC-1999  
LOCUS wg96h11.x1 NCI\_CGAP\_Kid11 Homo sapiens cDNA clone IMAGE:2379141 3',  
DEFINITION mRNA sequence.  
ACCESSION AI910763  
VERSION AI910763.1 GI:5630499  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.  
REFERENCE 1 (bases 1 to 302)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
CONTACT: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www.bio.lnl.gov/dbfp/image/image.html](http://www.bio.lnl.gov/dbfp/image/image.html)  
Insert Length: 1427 Std Error: 0.00  
Seq primer: -400p from Gibco  
High quality sequence scop: 297.  
Location/Qualifiers  
1..302  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2379141"  
/clone\_lib="NCI\_CGAP\_Kid11"  
/lab\_host="DH10B"  
/note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with  
a modified polylinker; Site: 1: Not I; Site 2: Eco RI;  
Plasmid DNA from the normalized library NCI\_CGAP\_Kid3 was  
prepared, and ss circles were made in vitro. Following HAP

purification, this DNA was used as tracer in a subtractive  
hybridization reaction. The driver was PCR-amplified cDNAs  
from a pool of 5,000 clones made from the same library  
(clonoids 132376-132911, 1456007-1456775, and  
1500552-1502855). Subtraction by Bento Soares and M.  
Fatima Bonaldo."

BASE COUNT 113 a 49 c 44 g 96 t

ORIGIN

Query Match 100.0%; Score 165; DB 9; Length 302;  
Best Local Similarity 100.0%; Pred. No. 1.3e-21;  
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCTTGAGTAAGTATTTTGGAGAGTTCATAGCAATGGAACATACATAAATGA 60  
|||||  
DB 4 CTCTTGAGTAAGTATTTTGGAGAGTTCATAGCAATGGAACATACATAAATGA 63  
QY 61 CACACCACTGTTGACAAATGAAAAAAGACATTTGATATTTTCACCTTTTAACTT 120  
|||||  
DB 64 CACACCACTGTTGACAAATGAAAAAAGACATTTGATATTTTCACCTTTTAACTT 123  
QY 121 AAAAAATGATTCAGTTAAACAAACAAAGTTGATTTTAACTT 165  
|||||  
DB 124 AAAAAATGATTCAGTTAAACAAACAAAGTTGATTTTAACTT 168

RESULT 5  
AI969490 303 bp mRNA linear EST 20-OCT-2000  
LOCUS wz67f01.x1 NCI\_CGAP\_Mel15 Homo sapiens cDNA clone IMAGE:2563125 3',  
DEFINITION mRNA sequence.  
ACCESSION AI969490  
VERSION AI969490.1 GI:5766308  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.  
REFERENCE 1 (bases 1 to 303)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
CONTACT: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Chris Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D. CDNA Library Preparation: Life  
Technologies, Inc. CDNA Library Arrayed by: Christa Prange, The  
I.M.A.G.E. Consortium Dyt Sequencing by: Washington University  
Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www.bio.lnl.gov/dbfp/image/image.html](http://www.bio.lnl.gov/dbfp/image/image.html)  
Insert Length: 449 Std Error: 0.00  
Seq primer: -400p from Gibco.  
Location/Qualifiers  
1..303  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2563129"  
/clone\_lib="NCI\_CGAP\_Mel15"  
/tissue\_type="malignant melanoma, metastatic to lymph  
node"  
/lab\_host="DH10B"  
/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: SalI;  
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies."

BASE COUNT 114 a 49 c 44 g 96 t

ORIGIN

Query Match 100.0%; Score 165; DB 9; Length 303;  
Best Local Similarity 100.0%; Pred. No. 1.3e-21;  
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CTCCTTGAAGTACTTATTTTGGAGAGTTCATAGCATTAAGACATACATAAATGA 60
DB 4 CTCCTTGAAGTACTTATTTTGGAGAGTTCATAGCATTAAGACATACATAAATGA 63
QY 61 CACACCACTGTTGACATGAAAAAAGAGCATTGATTTTCCAGCTTTTAAGTT 120
DB 64 CACACCACTGTTGACATGAAAAAAGAGCATTGATTTTCCAGCTTTTAAGTT 123
QY 121 AAAAAATGATTCAGTTAAACAAAAGTTAGATATTTTAG 165
DB 124 AAAAAATGATTCAGTTAAACAAAAGTTAGATATTTTAG 168

RESULT 6
LOCUS AA043474 381 bp mRNA linear EST 30-DEC-1998
DEFINITION q07f12.x1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone
IMAGE:1940207 3', mRNA sequence.
ACCESSION A1351615
VERSION A1351615
KEYWORDS EST.
ORGANISM Homo sapiens
SOURCE human.
REFERENCE 1 (bases 1 to 356)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL Contact: Robert Strassberg, Ph.D.
Email: c9apbs-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Glibco
High quality sequence stop: 310.
FEATURES
location/Qualifiers
1..356
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1940207"
/clone_lib="Soares_total_fetus_Nb2HF8_9w"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
/note="Vector: p7773D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - oligo(dT) primer (5'
TGTACCAATCTGAAGTGGAGCGCGCTTAATTTTCTTTTCTTTT 3').
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p7773 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 130 a 53 c 51 g 122 t
ORIGIN
Query Match 100.0%; Score 165; DB 9; Length 356;
Best Local Similarity 100.0%; Pred. No. 1.2e-21;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTCCTTGAAGTACTTATTTTGGAGAGTTCATAGCATTAAGACATACATAAATGA 60
DB 12 CTCCTTGAAGTACTTATTTTGGAGAGTTCATAGCATTAAGACATACATAAATGA 71
QY 61 CACACCACTGTTGACATGAAAAAAGAGCATTGATTTTCCAGCTTTTAAGTT 120
DB 72 CACACCACTGTTGACATGAAAAAAGAGCATTGATTTTCCAGCTTTTAAGTT 131
QY 121 AAAAAATGATTCAGTTAAACAAAAGTTAGATATTTTAG 165
DB 132 AAAAAATGATTCAGTTAAACAAAAGTTAGATATTTTAG 176

```

```

RESULT 7
LOCUS AA043474 381 bp mRNA linear EST 04-SEP-1996
DEFINITION zK54e09.s1 Soares_pregnant_uterus_NbHPU Homo sapiens cDNA clone
IMAGE:486664 3', mRNA sequence.
ACCESSION AA043474
VERSION AA043474.1 GI:1521330
KEYWORDS EST.
ORGANISM Homo sapiens
SOURCE human.
REFERENCE 1 (bases 1 to 381)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Ritkin,L., Rohlfing,T., Soares,M., Tan,F., Trevisakis,E., Waterston
R., Williamson,A., Wohlmann,P. and Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
JOURNAL Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40M13 fwd. from Amersham
High quality sequence stop: 301.
FEATURES
location/Qualifiers
1..381
/organism="Homo sapiens"
/db_xref="GDB:3760378"
/db_xref="taxon:9606"
/clone="IMAGE:486664"
/clone_lib="Soares_pregnant_uterus_NbHPU"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: uterus; Vector: p7773-Pac; Site_1: Not I;
Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer (5'
AAGTGAAGATTCGCGCGCCCTTTTCTTTTCTTTTCTTTT 3'),
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p7773 vector. Library
went through one round of normalization. Library
constructed by M. Fatima Bonaldo."
BASE COUNT 142 a 59 c 59 g 121 t
ORIGIN
Query Match 100.0%; Score 165; DB 9; Length 381;
Best Local Similarity 100.0%; Pred. No. 1.1e-21;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTCCTTGAAGTACTTATTTTGGAGAGTTCATAGCATTAAGACATACATAAATGA 60
DB 1 CTCCTTGAAGTACTTATTTTGGAGAGTTCATAGCATTAAGACATACATAAATGA 60
QY 61 CACACCACTGTTGACATGAAAAAAGAGCATTGATTTTCCAGCTTTTAAGTT 120
DB 61 CACACCACTGTTGACATGAAAAAAGAGCATTGATTTTCCAGCTTTTAAGTT 120
QY 121 AAAAAATGATTCAGTTAAACAAAAGTTAGATATTTTAG 165
DB 121 AAAAAATGATTCAGTTAAACAAAAGTTAGATATTTTAG 165

RESULT 8
LOCUS H08164 412 bp mRNA linear EST 23-JUN-1995

```

DEFINITION y187e01.s1 Soares infant brain INIB Homo sapiens cDNA clone  
IMAGE:44945 3', mRNA sequence.  
ACCESSION H08164  
VERSION H08164.1 GI:872986  
KEYWORDS EST  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 412)  
AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Harkins, M., Holman,  
M., Holman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,  
Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisakis, E., Waterston,  
R., Williamson, A., Wohlmann, P. and Wilson, R.  
The WashU-Merck EST Project  
Unpublished (1995)  
CONTACT: Wilson RK  
JOURNAL  
TITLE Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
High quality sequence stops: 337  
Source: IMAGE Consortium, LLNL  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: Promega -21ml3  
High quality sequence stop: 337.  
FEATURES  
source  
1. 412  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="Gene:417486"  
/db\_xref="taxon:9606"  
/clone="IMAGE:44945"  
/clone.lib="Soares infant brain INIB"  
/sex="female"  
/dev\_stage="73 days post natal"  
/lab\_host="DH10B (ampicillin resistant)"  
/note="Organ: whole brain; Vector: latmid BA; Site: 1: Not  
I; Site: 2: Hind III; 1st strand cDNA was primed with a Not  
I - oligo(dT) primer [5',  
AACTGACAAATTCGCCGCGAGCAATTTTCTTTTCTTTT 3'];  
double-stranded cDNA was ligated to Hind III adaptors  
(Pharmacia), digested with Not I and directionally cloned  
into the Not I and Hind III sites of the latmid BA vector.  
Library went through one round of normalization. Library  
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 143 a 65 c 64 g 139 t 1 others  
ORIGIN

Query Match 100.0%; Score 165; DB 14; Length 412;  
Best Local Similarity 100.0%; Pred. No. 1.1e-21;  
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCCTTGAGTACTTATTTGGAGAGTTCATAGCAATGGAACATACATAAATGA 60  
|||||  
DB 14 CTCCTTGAGTACTTATTTGGAGAGTTCATAGCAATGGAACATACATAAATGA 73  
QY 61 CACACACACTGTTGACATGAAAAAAGACATTTGATTTTCCAGCTTTTAAGTT 120  
|||||  
DB 74 CACACACACTGTTGACATGAAAAAAGACATTTGATTTTCCAGCTTTTAAGTT 133  
QY 121 AAAAAATGATTCAGTTAAAAAAGTTAGATTTTAAAGTTAG 165  
|||||  
DB 134 AAAAAATGATTCAGTTAAAAAAGTTAGATTTTAAAGTTAG 178

RESULT 9  
LOCUS A1049699 423 bp mRNA linear EST 09-JUL-1998  
DEFINITION an33d09.x1 Gessler Wlms tumor Homo sapiens cDNA clone  
IMAGE:1700465 3', mRNA sequence.  
ACCESSION A1049699

VERSION A1049699.1 GI:3298816  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 423)  
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,  
Kriman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin,  
J., Moore, B., Schellenberg, K., Sepcece, M., Tan, F., Theising, B.,  
White, X., Wylie, T., Waterston, R. and Wilson, R.  
WashU-NCI human EST Project  
Unpublished (1997)  
CONTACT: Wilson RK  
JOURNAL  
TITLE Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -40ml3 fwd. ET from AmerSham.  
FEATURES  
source  
1. 423  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1700465"  
/clone.lib="Gessler Wlms tumor"  
/sex="pooled (6)"  
/lab\_host="DH10B"  
/note="Vector: pSPOR1; Site: 1: SalI; Site: 2: NotI; RNA  
was prepared from a pool of 6 anonymous Wlms' tumor RNAs.  
RNA was prepared by acid-phenol, followed by one round of  
oligo dt selection. cDNA library preparation was with  
the BRL/Life Tech. Superscript Plasmid system. An  
oligo dT NotI primer for first strand synthesis generated  
ggcgccgccct)n at the 3' end of the clones. A 5' SalI  
adaptor was used with sequence 5'-gtcagccacgcgcgcg-3'.  
Resulting cDNAs were size selected (average size 2 kb),  
NotI digested, and ligated into NotI/SalI-cut pSPOR1.  
Library was constructed by Dr. Manfred Gessler."

BASE COUNT 155 a 65 c 60 g 143 t  
ORIGIN

Query Match 100.0%; Score 165; DB 9; Length 423;  
Best Local Similarity 100.0%; Pred. No. 1.1e-21;  
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCCTTGAGTACTTATTTGGAGAGTTCATAGCAATGGAACATACATAAATGA 60  
|||||  
DB 14 CTCCTTGAGTACTTATTTGGAGAGTTCATAGCAATGGAACATACATAAATGA 73  
QY 61 CACACACACTGTTGACATGAAAAAAGACATTTGATTTTCCAGCTTTTAAGTT 120  
|||||  
DB 74 CACACACACTGTTGACATGAAAAAAGACATTTGATTTTCCAGCTTTTAAGTT 133  
QY 121 AAAAAATGATTCAGTTAAAAAAGTTAGATTTTAAAGTTAG 165  
|||||  
DB 134 AAAAAATGATTCAGTTAAAAAAGTTAGATTTTAAAGTTAG 178

RESULT 10  
LOCUS AW592865 423 bp mRNA linear EST 22-MAR-2000  
DEFINITION hg04b07.x1 Soares\_NFL\_T.GBC.S1 Homo sapiens cDNA clone  
IMAGE:2944597 3', mRNA sequence.  
ACCESSION AW592865  
VERSION AW592865.1 GI:7280057  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE	1 (bases 1 to 423)
AUTHORS	NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a> This clone is available royalty-free through LNL; contact the IMAGE Consortium ( <a href="mailto:info@image.llnl.gov">info@image.llnl.gov</a> ) for further information. Seq primer: -400p from Glibdo High quality sequence stopers 209. location/Qualifiers
FEATURES	
source	1..423 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:2944597" /clone_lib="Soares_NFL_T_GBC_S1" /lab_host="DH10B" /note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not 1; Site_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NBHL19W, testis NHT, and B-cell NCI-CGAP-GCB1) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726008-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo."
BASE COUNT	157 a 66 c 63 g 137 t
ORIGIN	
Query Match	100.0%; Score 165; DB 10; Length 423;
Best Local Similarity	100.0%; Pred. No. 1 ie-21;
Matches 165;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 CTCTTTGAGTACTTTATTTTGGAGAGACTTCATPAAGCATTAGAGAACATACATAAAATGA 60       6 CTCTTTAGTAACTTTATTTTGGAGAGACTTCATPAAGCATTAGAGAACATACATAAAATGA 65
Db	
QY	61 CACACCACTGTTGCAATGAAACAAAAACAGCATTTGCAATTTTCCAGCTTTTAACTT 120       66 CACACCACTGTTGCAATGAAACAAAAACAGCATTTGATATTTTCCAGCTTTTAACTT 125
Db	
QY	121 AAAAAATGATTCAGTTAAACAAAAACAAAGTTAGATTTTGTAG 165       126 AAAAAATGATTCAGTTAAACAAAAACAAAGTTAGATTTTGTAG 170
Db	
RESULT 11	
LOCUS	A1041596 444 bp mRNA linear EST 24-SEP-1998
DEFINITION	ox56h06.x1 Soares_NHMPu_S1 Homo sapiens cDNA clone IMAGE:1560579
LOCUS	3', mRNA sequence.
ACCESSION	A1041596
VERSION	A1041596.1 GI:3280790
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
COMMENT	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 444) NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> . National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a> This clone is available royalty-free through LNL; contact the IMAGE Consortium ( <a href="mailto:info@image.llnl.gov">info@image.llnl.gov</a> ) for further information. Insert length: 827 Std Error: 0.00 Seq primer: -40ml3 fwd. ET from Amersham

FEATURES	High quality sequence stop: 414.
Source	Location/Qualifiers
	1. 444
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="IMAGE:1660379"
	/clone_lib="Soares_NbHMPu_S1"
	/tissue_type="Pooled human melanocyte, fetal heart, and pregnant uterus"
	/lab_host="DH10B"
	/note="Organ: mixed (see below); Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not i; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NbHM, pregnant uterus NbHPU, and fetal heart NbH19W) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."
BASE COUNT	166 a 69 c 71 g 138 t
ORIGIN	
Query Match	100.0%; Score 165; DB 9; Length 444;
Best Local Similarity	100.0%; Pred. No. 1.1e-21;
Matches 165; Conservative	0; Mismatches 0; Gaps 0;
Oy	1 CTCTTGGAGTACTTTATTTTGGAGAGTTCATAGCATAGGAACATACATAAATGA 60
Db	4 CTCTTGGAGTACTTTATTTTGGAGAGTTCATAGCATAGGAACATACATAAATGA 63
Oy	61 CACACCACTGTGCACATGAAAAAACAACGACATTGATATTTCCAGCTTTTAAAGT 120
Db	64 CACACCACTGTGCACATGAAAAAACAACGACATTGATATTTCCAGCTTTTAAAGT 123
Oy	121 AAAAAATGATTCAGTTAAACCAAAACAAAGTTAGATATTTTG 165
Db	124 AAAAAATGATTCAGTTAAACCAAAACAAAGTTAGATATTTTG 168
RESULT 12	
AM072470	
LOCUS	450 bp mRNA linear EST 20-OCT-2000
DEFINITION	xa06d07.x1 Soares_NFL_T-GBC_S1 Homo sapiens CDNA clone.
IMAGE:2567533 3	, mRNA sequence.
AM072470	
AM072470.1	GI:6027468
EST.	
SOURCE	human.
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 450)
AUTHORS	NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaps-remail.nih.gov This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Insert Length: 1130 Std Error: 0.00 Seq primer: -40UP from Gibco High quality sequence stop: 447. Location/Qualifiers
FEATURES	
Source	1. 450
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="IMAGE:2567533"
	/clone_lib="Soares_NFL_T-GBC_S1"
	/lab_host="DH10B"
	/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with

a modified polylinker. Site\_1: Not I; Site\_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NBHL19W, testis NHT, and B-cell NCI-CGAP\_GCB1) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of 1.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 170 a 70 c 140 t

Query Match 100.0%; Score 165; DB 10; Length 450;  
Best Local Similarity 100.0%; Pred. No. 1.1e-21;  
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCCTTGAGTAAGTATTTTGGAGAGTTCATAGCATTAAGCAATCAATGA 60  
|||||  
DB 4 CTCCTTGAGTAAGTATTTTGGAGAGTTCATAGCATTAAGCAATCAATGA 63  
61 CACACCACTGTTGACAAATGAAAAAAGACATTTGATATTTCCAGCTTTTAAGTT 120  
64 CACACCACTGTTGACAAATGAAAAAAGACATTTGATATTTCCAGCTTTTAAGTT 123  
QY 121 AAAAAATGATTCAGTTAAACAAACAAAGTTTATGATATTTTAA 165  
|||||  
DB 124 AAAAAATGATTCAGTTAAACAAACAAAGTTTATGATATTTTAA 168

RESULT 13  
A1573107 451 bp mRNA linear EST 14-MAY-1999  
LOCUS uc69f01.x1 NCI\_CGAP\_Pani Homo sapiens cDNA clone IMAGE:2223577 3',  
DEFINITION mRNA sequence.  
ACCESSION A1573107  
VERSION A1573107.1 GI:4536481  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE 1 (bases 1 to 451)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
JOURNAL COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Life Technologies catalog #: 11548-013  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www.biol.llnl.gov/bbrp/image/image.html](http://www.biol.llnl.gov/bbrp/image/image.html)  
Insert Length: 3468 Std Error: 0.00  
Seq primer: -400P from Gibco  
High quality sequence stop: 405  
POLYA=NO.

FEATURES  
SOURCE Location/Qualifiers  
1..451  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2223577"  
/clone\_lib="NCI\_CGAP\_Pani"  
/tissue\_type="adenocarcinoma"  
/lab\_host="DH10B"  
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.72 kb. Life Technologies catalog #:  
11548-013"

BASE COUNT 169 a 71 c 70 g 141 t

ORIGIN

Query Match 100.0%; Score 165; DB 9; Length 451;  
Best Local Similarity 100.0%; Pred. No. 1.1e-21;  
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCCTTGAGTAAGTATTTTGGAGAGTTCATAGCATTAAGCAATCAATGA 60  
|||||  
DB 9 CTCCTTGAGTAAGTATTTTGGAGAGTTCATAGCATTAAGCAATCAATGA 68  
61 CACACCACTGTTGACAAATGAAAAAAGACATTTGATATTTCCAGCTTTTAAGTT 120  
69 CACACCACTGTTGACAAATGAAAAAAGACATTTGATATTTCCAGCTTTTAAGTT 128  
QY 121 AAAAAATGATTCAGTTAAACAAACAAAGTTTATGATATTTTAA 165  
|||||  
DB 129 AAAAAATGATTCAGTTAAACAAACAAAGTTTATGATATTTTAA 173

RESULT 14  
AW770384 459 bp mRNA linear EST 04-MAY-2000  
LOCUS h180d09.x1 NCI\_CGAP\_Kid13 Homo sapiens cDNA clone IMAGE:3007505 3',  
DEFINITION mRNA sequence.  
ACCESSION AW770384  
VERSION AW770384.1 GI:7702426  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE 1 (bases 1 to 459)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
JOURNAL COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Chris Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life  
Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The  
I.M.A.G.E. Consortium DNA Sequencing by: Washington University  
Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL, send email to:  
[info@image.llnl.gov](mailto:info@image.llnl.gov)  
Seq primer: -400P from Gibco  
High quality sequence stop: 415.

FEATURES  
SOURCE Location/Qualifiers  
1..459  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3007505"  
/clone\_lib="NCI\_CGAP\_Kid13"  
/tissue\_type="2 pooled Wilms' tumors, one primary and one  
metastatic to brain"  
/lab\_host="DH10B"  
/note="Organ: kidney; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies."

BASE COUNT 173 a 71 c 73 g 142 t

ORIGIN

Query Match 100.0%; Score 165; DB 10; Length 459;  
Best Local Similarity 100.0%; Pred. No. 1.1e-21;  
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCCTTGAGTAAGTATTTTGGAGAGTTCATAGCATTAAGCAATCAATGA 60  
|||||  
DB 6 CTCCTTGAGTAAGTATTTTGGAGAGTTCATAGCATTAAGCAATCAATGA 65  
61 CACACCACTGTTGACAAATGAAAAAAGACATTTGATATTTCCAGCTTTTAAGTT 120  
66 CACACCACTGTTGACAAATGAAAAAAGACATTTGATATTTCCAGCTTTTAAGTT 125

OY 121 AAAAAATGATTCAGTTAAACAAAAGTTAGATATTAG 165  
 Db 126 AAAAAATGATTCAGTTAAACAAAAGTTAGATATTAG 170

RESULT 15

AA479302 462 bp mRNA linear EST 08-AUG-1997  
 LOCUS zv21f08.s1 Soares\_NbHMPu\_S1 Homo sapiens cDNA clone IMAGE:754311  
 DEFINITION 3', mRNA sequence.

ACCESSION AA479302

VERSION AA479302.1 GI:2207858

KEYWORDS EST.

SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 462)  
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,  
 Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,  
 Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie,  
 T., Waterston, R. and Wilson, R.  
 Mashu-Merck EST Project 1997

TITLE JOURNAL  
 COMMENT Unpublished (1997)  
 Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Seq primer: -41m13 fwd. ET from Amersham  
 High quality sequence stop: 456.  
 Location/Qualifiers

FEATURES

source

1..462  
 /organism="Homo sapiens"  
 /db\_xref="GDB:5977197"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:754311"  
 /clone\_id="Soares\_NbHMPu\_S1"  
 /tissue\_type="Pooled human melanocyte, fetal heart, and  
 pregnant uterus"  
 /lab\_host="DH10B"  
 /note="Organ: mixed (see below); Vector: pT7T3D-Pac  
 (Pharmacia) with a modified polylinker; Site\_1: Not I;  
 Site\_2: Eco RI; Equal amounts of plasmid DNA from three  
 normalized libraries (melanocyte 2NbHM, pregnant uterus  
 NbHPU, and fetal heart NbH19W) were mixed, and ss circles  
 were made in vitro. Following HAP purification, this DNA  
 was used as tracer in a subtractive hybridization  
 reaction. The driver was PCR-amplified cDNAs from pools of  
 5,000 clones made from the same 3 libraries. The pools  
 consisted of I.M.A.G.E. clones 260232-265223,  
 340488-345479, and 484488-489479."

BASE COUNT 176 a 70 c 75 g 141 t  
 ORIGIN

Query Match 100.0%; Score 165; DB 9; Length 462;  
 Best Local Similarity 100.0%; Pred. No. 1,1e-21;  
 Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CTCTTTGAGTAACTTTATTTGGACGAGTTCATTAAGCATTTAGAACATATAAATGA 60  
 Db 1 CTCTTTGAGTAACTTTATTTGGAGAGTTCATTAAGCATTTAGAACATATAAATGA 60  
 OY 61 CACACCACTGTGACATGAAAAAACAAGCATTTGATTTTCCAGCTTTTAAGTT 120  
 Db 61 CACACCACTGTGACATGAAAAAACAAGCATTTGATTTTCCAGCTTTTAAGTT 120  
 OY 121 AAAAAATGATTCAGTTAAACAAAAGTTAGATATTAG 165  
 Db 121 AAAAAATGATTCAGTTAAACAAAAGTTAGATATTAG 165

Db 121 AAAAAATGATTCAGTTAAACAAAAGTTAGATATTAG 165  
 Search completed: June 21, 2003, 03:38:08  
 Job time : 496.851 secs

## ALIGNMENTS

for SEQ ID NO: 390

RESULT 1  
ABL67356  
ID ABL67356 standard; DNA; 386 BP.  
XX  
AC ABL67356;  
XX  
DT 15-MAY-2002 (first entry)  
XX  
DE Thyroid cancer related gene sequence SEQ ID NO:5693.  
XX  
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200194629-A2.  
XX  
PD 13-DEC-2001.  
XX  
PF 30-MAY-2001; 2001WO-US10838.  
XX  
PR 05-JUN-2000; 2000US-209473P.  
PR 05-JUN-2000; 2000US-209531P.  
PR 18-SEP-2000; 2000US-233133P.  
PR 18-SEP-2000; 2000US-233617P.  
PR 20-SEP-2000; 2000US-234009P.  
PR 20-SEP-2000; 2000US-234034P.  
PR 20-SEP-2000; 2000US-234052P.  
PR 22-SEP-2000; 2000US-234509P.  
PR 22-SEP-2000; 2000US-234567P.

PR 25-SEP-2000; 2000US-234923P.  
PR 25-SEP-2000; 2000US-234924P.  
PR 25-SEP-2000; 2000US-235077P.  
PR 25-SEP-2000; 2000US-235082P.  
PR 25-SEP-2000; 2000US-235134P.  
PR 25-SEP-2000; 2000US-235280P.  
PR 26-SEP-2000; 2000US-235637P.  
PR 26-SEP-2000; 2000US-235638P.  
PR 27-SEP-2000; 2000US-235711P.  
PR 27-SEP-2000; 2000US-235720P.  
PR 27-SEP-2000; 2000US-235840P.  
PR 28-SEP-2000; 2000US-235863P.  
PR 28-SEP-2000; 2000US-236028P.  
PR 28-SEP-2000; 2000US-236032P.  
PR 28-SEP-2000; 2000US-236033P.  
PR 28-SEP-2000; 2000US-236034P.  
PR 28-SEP-2000; 2000US-236109P.  
PR 29-SEP-2000; 2000US-236111P.  
PR 29-SEP-2000; 2000US-236842P.  
PR 29-SEP-2000; 2000US-236843P.  
PR 02-OCT-2000; 2000US-237172P.  
PR 02-OCT-2000; 2000US-237173P.  
PR 02-OCT-2000; 2000US-237278P.  
PR 02-OCT-2000; 2000US-237294P.  
PR 02-OCT-2000; 2000US-237295P.  
PR 02-OCT-2000; 2000US-237316P.  
PR 03-OCT-2000; 2000US-237425P.  
PR 03-OCT-2000; 2000US-237598P.  
PR 03-OCT-2000; 2000US-237604P.  
PR 03-OCT-2000; 2000US-237606P.  
PR 03-OCT-2000; 2000US-237608P.  
PR 01-NOV-2000; 2000US-244867P.  
PR 01-NOV-2000; 2000US-245084P.  
XX  
XX (AVAL-) AVALON PHARM.  
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
PI Soppet DR, Weaver Z;  
XX  
XX WPI: 2002-188264/24.

Screening for anti-neoplastic agent involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, and determining a change in expression of a gene of a signature gene set -  
Claim 1: SEQ ID 5693: 44pp: English.

The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (s) selected from 847 sequences (given in ABL6164 to ABL70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophagel, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms' tumour.

Sequence 386 BP: 142 A: 56 C: 63 G: 122 T: 3 other:

Query Match 99.2%; Score 383; DB 24; Length 386;  
Best Local Similarity 100.0%; Pred. No. 5, 8e-68;  
Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 AAATTATATATACATGTTTATTAGAGACACACTTTATGTAAATTTACATTATGCA 60  
QY 61 AAAAAATCAAAAATTTTACAAAATCTTGAAGACAGATGTCATGTTCTAATTACAAAT 120  
DB 61 AAAAAATCAAAAATTTTACAAAATCTTGAAGACAGATGTCATGTTCTAATTACAAAT 120  
QY 121 CCAAAGTAGTAATACATGCTTTAAAGTCACATTTATAGAGTTGTGTTTACAAAT 180  
DB 121 CCAAAGTAGTAATACATGCTTTAAAGTCACATTTATAGAGTTGTGTTTACAAAT 180  
QY 181 CTGGTTAAAGAGGAGCTACAAAGTTATACATATATAGCAAGACAGCTTGCTA 240  
DB 181 CTGGTTAAAGAGGAGCTACAAAGTTATACATATATAGCAAGACAGCTTGCTA 240  
QY 241 GGCTACATTTCCCATTTGAAATCTAGTGGTCTCTTTTACACCATTTAGGGGATTTTAA 300  
DB 241 GGCTACATTTCCCATTTGAAATCTAGTGGTCTCTTTTACACCATTTAGGGGATTTTAA 300  
QY 301 ATGGGNNAAAAAATCAATATTAACATATGCGCTTCAAAATTTGTAACCTGTACCC 360  
DB 301 ATGGGNNAAAAAATCAATATTAACATATGCGCTTCAAAATTTGTAACCTGTACCC 360  
QY 361 NATACTGGGNNATGAGAGGCTGTGG 386  
DB 361 NATACTGGGNNATGAGAGGCTGTGG 386

1 AAATTATATATACATGTTTATTAGAGACACACTTTATGTAAATTTACATTATGCA 60  
|||||

Mon Jun 23 10:00:41 2003

us-09-96

RESULT 1

LOCUS

DEFINITION

R27957 386 bp mRNA linear EST 25-APR-1995  
 yhs7b02.s1 Soares placenta Nb2HP Homo sapiens cDNA clone  
 IMAGE:133803 3', mRNA sequence.

ACCESSION

VERSION

R27957  
 R27957.1 GI:784092

KEYWORDS

SOURCE

human.  
 Homo sapiens

REFERENCE

Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Eukaryota; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

AUTHORS

Hallier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman  
 M., Hulman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,  
 Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston  
 R., Williamson, A., Wohldmann, P. and Wilson, R.

TITLE

The Washu-Merck EST Project  
 Unpublished (1995)

JOURNAL

Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800  
 Fax: 314 286 1810

Email: est@watson.wustl.edu

Insert Size: 820

High quality sequence stops: 302

Source: IMAGE Consortium, LNL

This clone is available royalty-free through LNL; contact the

FEATURES  
 source  
 Location/Qualifiers  
 1..386  
 /organism="Homo sapiens"  
 /db\_xref="GDB:539557"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:133803"  
 /clone\_11b="Soares placenta Nb2HP"  
 /sex="Female"  
 /dev\_stage="placenta obtained at birth (full term)"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /note="Organ: Placenta; Vector: pT73D (Pharmacia) with a  
 modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st  
 strand cDNA was primed with a Not I - oligo(dT) primer [5'  
 AACTGGAAGATTGCGCGCCGACAGATTGTTTCTTTTCTTTT 3'],  
 double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pT73 vector. Library  
 went through one round of normalization. Library  
 constructed by Bento Soares and M.Fatima Bonafide."

BASE COUNT 142 a 56 c 63 g 122 t 3 others

Query Match

Best Local Similarity 100.0%; Pred. No. 4.4e-65;

Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	AAATTATATATACATGTTTATTAAGGCAAACTTTATGTAATAATTACATTATGA	60
DB	1	AAATTATATATACATGTTTATTAAGGCAAACTTTATGTAATAATTACATTATGA	60
QY	61	AAAAATCAAAAATATTACAAAATCTTGAAGACAGATGTCATTTCTTAATTCAAAT	120
DB	61	AAAAATCAAAAATATTACAAAATCTTGAAGACAGATGTCATTTCTTAATTCAAAT	120
QY	121	CCAAAGTAGTAATTAACATCCCTTTAAACCTCATTATTAAGTTGTTTACAAAT	180
DB	121	CCAAAGTAGTAATTAACATCCCTTTAAACCTCATTATTAAGTTGTTTACAAAT	180
QY	181	CTTGTTTAAAGAGGACAGTACAAAGTTTCACTATATTAAGCAAGAACAGCTTGCTA	240
DB	181	CTTGTTTAAAGAGGACAGTACAAAGTTTCACTATATTAAGCAAGAACAGCTTGCTA	240
QY	241	GGGTACATTTCCCATTTGAAAATCTACTGGGTCCTTTTACACATTAGGGGATTTTAA	300
DB	241	GGGTACATTTCCCATTTGAAAATCTACTGGGTCCTTTTACACATTAGGGGATTTTAA	300
QY	301	ATGGGNNAAAAAATCAATATTAACCTCATATGAGCTTCAAAAATGTGTAACTGTACCC	360
DB	301	ATGGGNNAAAAAATCAATATTAACCTCATATGAGCTTCAAAAATGTGTAACTGTACCC	360
QY	361	NATACCTGGGNNATGAGGCGCTGTGG 386	
DB	361	NATACCTGGGNNATGAGGCGCTGTGG 386	

Mon Jun 23 10:00:41 2003

us-09-964-

RESULT 1  
H05625 353 bp mRNA linear EST 21-JUN-1995  
LOCUS Y170b12.s1 Soares infant brain INIB Homo sapiens cDNA clone  
DEFINITION IMAGE:43318 3', mRNA sequence.  
ACCESSION H05625  
VERSION H05625.1 GI:869177  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 353)  
AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisan, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.  
TITLE The WashU-Merck EST Project  
JOURNAL Unpublished (1995)  
COMMENT Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
High quality sequence stops: 291  
Source: IMAGE Consortium, LLNL.  
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (infoimage.llnl.gov) for further information.

Mon Jun 23 10:00:41 2003

Seq primer: Promega -21m13  
High quality sequence stop: 310.  
Location/Qualifiers  
1..353  
/organism="Homo sapiens"  
/db\_xref="GDB:415859"  
/db\_xref="taxon:9606"  
/clone="IMAGE:43318"  
/clone\_11b="Soares infant brain INIB"  
/sex="female"  
/dev\_stage="73 days post natal"  
/lib\_host="DH10B (ampicillin resistant)"  
/note="Organ: whole brain; Vector: Latmid BA; Site\_1: Not I; Site\_2: Hind III; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'  
ACCTGCAAGAAATCGCGCGCGCAAGAAATTTTTTTTTTTTTTTT 3'];  
double-stranded cDNA was ligated to Hind III adaptors  
(Pharmacia), digested with Not I and directionally cloned  
into the Not I and Hind III sites of the Latmid BA vector.  
Library went through one round of normalization. Library  
constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 107 a 53 c 72 g 116 t 5 others  
ORIGIN

Query Match 98.6%; Score 348; DB 14; Length 353;  
Best Local Similarity 100.0%; Pred. No. 2.2e-43;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	TTTTTTTTTTTTTTTTTTTTTGTGCTCACAATAATGTCATTTTATGTGACACTAGTCACAAC	60
DB	1	TTTTTTTTTTTTTTTTTTTTTGTGCTCACAATAATGTCATTTTATGTGACACTAGTCACAAC	60
QY	61	AAATACATAATTTGCAAG	120
DB	61	AAATACATAATTTGCAAG	120
QY	121	GAACTGCAGTTTTCAGAGTACTACTATTTATTTAAAAAATACAAANAGAAAA	180
DB	121	GAACTGCAGTTTTCAGAGTACTACTATTTATTTAAAAAATACAAANAGAAAA	180
QY	181	ATGTATACACTACAGTAGGAGATTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	240
DB	181	ATGTATACACTACAGTAGGAGATTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	240
QY	241	GGGTTAAAAACATTTTCATGAGATTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG	300
DB	241	GGGTTAAAAACATTTTCATGAGATTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG	300
QY	301	ACCGTAATCTGGGCAACATCCGGGGCTTACCTTACGCTCTGGCAGCTGTGCG	353
DB	301	ACCGTAATCTGGGCAACATCCGGGGCTTACCTTACGCTCTGGCAGCTGTGCG	353

XX  
DT 15-MAY-2002 (first entry)  
XX  
DE Breast cancer related gene sequence SEQ ID NO:1584.  
XX  
Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KM stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KM cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
XX gene; ds.  
XX  
OS Homo sapiens.  
PN WO200194629-A2.  
PD 13-DEC-2001.  
XX  
PF 30-MAY-2001; 2001WO-US10838.  
XX  
PR 05-JUN-2000; 2000US-209473P.  
PR 05-JUN-2000; 2000US-209531P.  
PR 18-SEP-2000; 2000US-233133P.  
PR 18-SEP-2000; 2000US-233617P.  
PR 20-SEP-2000; 2000US-234009P.  
PR 20-SEP-2000; 2000US-234034P.  
PR 20-SEP-2000; 2000US-234052P.  
PR 22-SEP-2000; 2000US-234509P.  
PR 22-SEP-2000; 2000US-234567P.  
PR 25-SEP-2000; 2000US-234923P.  
PR 25-SEP-2000; 2000US-234924P.  
PR 25-SEP-2000; 2000US-235077P.  
PR 25-SEP-2000; 2000US-235082P.  
PR 25-SEP-2000; 2000US-235134P.  
PR 25-SEP-2000; 2000US-235280P.  
PR 26-SEP-2000; 2000US-235637P.  
PR 26-SEP-2000; 2000US-235638P.  
PR 27-SEP-2000; 2000US-235711P.  
PR 27-SEP-2000; 2000US-235720P.  
PR 27-SEP-2000; 2000US-235840P.  
PR 28-SEP-2000; 2000US-235863P.  
PR 28-SEP-2000; 2000US-236032P.  
PR 28-SEP-2000; 2000US-236033P.  
PR 28-SEP-2000; 2000US-236034P.  
PR 28-SEP-2000; 2000US-236109P.  
PR 28-SEP-2000; 2000US-236111P.  
PR 29-SEP-2000; 2000US-236842P.  
PR 29-SEP-2000; 2000US-236891P.  
PR 02-OCT-2000; 2000US-237172P.  
PR 02-OCT-2000; 2000US-237173P.  
PR 02-OCT-2000; 2000US-237278P.  
PR 02-OCT-2000; 2000US-237294P.  
PR 02-OCT-2000; 2000US-237295P.  
PR 02-OCT-2000; 2000US-237316P.  
PR 03-OCT-2000; 2000US-237425P.  
PR 03-OCT-2000; 2000US-237598P.  
PR 03-OCT-2000; 2000US-237604P.  
PR 03-OCT-2000; 2000US-237606P.  
PR 03-OCT-2000; 2000US-237608P.  
PR 01-NOV-2000; 2000US-244867P.  
PR 01-NOV-2000; 2000US-245084P.  
XX  
PA (AVAL-) AVALON PHARM.  
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horriqan S;  
PI Soppet DR, Weaver Z;  
XX  
DR WPI: 2002-188264/24.  
XX  
PT Screening for anti-neoplastic agent involves exposing cells to a  
PT chemical agent to be tested for anti-neoplastic activity, and  
PD determining a change in expression of a gene of a signature gene set

XX  
PS Claim 1; SEQ ID 1584; 44pp; English.  
XX  
CC The present invention describes a method (M1) for screening for an  
CC anti-neoplastic agent. The method involves exposing cells to a chemical  
CC agent to be tested for anti-neoplastic activity determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
CC to ABL70110), or is at least 95% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
CC activity and can be used in gene therapy. M1 can be used for screening  
CC an anti-neoplastic agent, and can be used for producing a product which  
CC is the data collected with respect to the anti-neoplastic agent as a  
CC result of M1, and the data is sufficient to convey the chemical  
CC structure and/or properties of the agent. M1 can be used in the  
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
CC carcinoma, papillary carcinoma and Wilm's tumour.  
XX  
SQ Sequence 353 BP; 107 A; 53 C; 72 G; 116 T; 5 other;  
Query Match 98.6%; Score 348; DB 24; Length 353;  
Best Local Similarity 100.0%; Pred. No. 1.6e-68;  
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTTT TTTT TTTT TTTT TTTT GCTCACAATGCTATTTATTTAGACTAGTGACACT 60  
DB 1 TTTT TTTT TTTT TTTT TTTT TTTT GCTCACAATGCTATTTATTTAGACTAGTGACACT 60  
QY 61 AATACATTAATTCAG 120  
DB 61 AATACATTAATTCAG 120  
QY 121 GAATGCACTGTTTTCAG 180  
DB 121 GAATGCACTGTTTTCAG 180  
QY 181 ATGTTATCACTACAGTAG 240  
DB 181 ATGTTATCACTACAGTAG 240  
QY 241 GGGTTAAACATTTATGAG 300  
DB 241 GGGTTAAACATTTATGAG 300  
QY 301 ACCGTAATCTGGGCAACATCCGGGGCTTACCTGAGCTCGGCACTGTGGC 353  
DB 301 ACCGTAATCTGGGCAACATCCGGGGCTTACCTGAGCTCGGCACTGTGGC 353  
RESULT 4  
ABL67431  
ID ABL67431 standard; DNA: 353 BP.  
XX  
AC ABL67431;  
XX  
DT 15-MAY-2002 (first entry)  
XX  
DE Thyroid cancer related gene sequence SEQ ID NO:5768.  
XX  
Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KM stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KM cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
XX gene; ds.  
XX  
OS Homo sapiens.  
PN WO200194629-A2.  
PD 13-DEC-2001.  
XX

PF	30-MAY-2001;	2001MO-US10838.
XX		
PR	05-JUN-2000;	2000US-209473P.
PR	05-JUN-2000;	2000US-209531P.
PR	18-SEP-2000;	2000US-231133P.
PR	18-SEP-2000;	2000US-233617P.
PR	20-SEP-2000;	2000US-234009P.
PR	20-SEP-2000;	2000US-234034P.
PR	20-SEP-2000;	2000US-234052P.
PR	22-SEP-2000;	2000US-234505P.
PR	22-SEP-2000;	2000US-234567P.
PR	25-SEP-2000;	2000US-234923P.
PR	25-SEP-2000;	2000US-234924P.
PR	25-SEP-2000;	2000US-235077P.
PR	25-SEP-2000;	2000US-235082P.
PR	25-SEP-2000;	2000US-235134P.
PR	25-SEP-2000;	2000US-235280P.
PR	26-SEP-2000;	2000US-235637P.
PR	26-SEP-2000;	2000US-235638P.
PR	27-SEP-2000;	2000US-235711P.
PR	27-SEP-2000;	2000US-235720P.
PR	27-SEP-2000;	2000US-235840P.
PR	27-SEP-2000;	2000US-235863P.
PR	28-SEP-2000;	2000US-236028P.
PR	28-SEP-2000;	2000US-236032P.
PR	28-SEP-2000;	2000US-236033P.
PR	28-SEP-2000;	2000US-236034P.
PR	28-SEP-2000;	2000US-236109P.
PR	28-SEP-2000;	2000US-236111P.
PR	29-SEP-2000;	2000US-236842P.
PR	29-SEP-2000;	2000US-236891P.
PR	02-OCT-2000;	2000US-237172P.
PR	02-OCT-2000;	2000US-237173P.
PR	02-OCT-2000;	2000US-237278P.
PR	02-OCT-2000;	2000US-237294P.
PR	02-OCT-2000;	2000US-237295P.
PR	02-OCT-2000;	2000US-237316P.
PR	03-OCT-2000;	2000US-237425P.
PR	03-OCT-2000;	2000US-237598P.
PR	03-OCT-2000;	2000US-237604P.
PR	03-OCT-2000;	2000US-237606P.
PR	03-OCT-2000;	2000US-237608P.
PR	01-NOV-2000;	2000US-244867P.
PR	01-NOV-2000;	2000US-245084P.

PT Screening for anti-neoplastic agent involves exposing cells to a  
PT chemical agent to be tested for anti-neoplastic activity, and  
PT determining a change in expression of a gene of a signature gene set -  
XX  
PS Claim 1, SEQ ID 5768; 44pp; English.

CC The present invention describes a method (M1) for screening for an  
CC anti-neoplastic agent. The method involves exposing cells to a chemical  
CC agent to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 8447 sequences (given in AB161664  
CC to AB170110), or is at least 95% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
CC activity and can be used in gene therapy. M1 can be used for screening  
CC an anti-neoplastic agent, and can be used for producing a product which  
CC is the data collected with respect to the anti-neoplastic agent as a  
CC result of M1, and the data is sufficient to convey the chemical  
CC structure and/or properties of the agent. M1 can be used in the  
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
CC lymphoma, carcinoma, clear cell cancer, infiltrating ductal cancer

CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's tumour.  
CC carcinoma, papillary carcinoma and Wilm's tumour.  
XX  
SQ Sequence 353 BP; 107 A; 53 C; 72 G; 116 T; 5 other;

Query Match	98.6%;	Score 348;	DB 24;	Length 353;
Best Local Similarity	100.0%;	Pred. No. 1.6e-68;		
Matches 353;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

QY	1	TTTTTTTTTTTTTTTTTTTTTTTTTTCTCCACAAAGTCAATTTTATGACACGTGTGACACAAC	60
Db	1	TTTTTTTTTTTTTTTTTTTTTTTGGCTCCACAAAGTCAATTTTATGACACGTGTGACACAAC	60
QY	61	AAATACATATATGCAAGAGAGTGAACGTGTTCAACAGAAATGTCACATAGAGTTA	120
Db	61	AAATACAAATATGGAAGAGAGTGAACCGTGTCAACAGAAATGTCACATAGAGTTA	120
QY	121	GAACGTGCACTTTTTCGAAGTACTGACTATATTATTTAAAAAAAATACAAANAGAAAA	180
Db	121	GAACGTGCACTTTTTCGAAGTACTGACTATATTATTTAAAAAAAATACCAANAGAAAA	180
QY	181	ATGTATATCCTACAAAGTAGGATTTAGGAAGAGAAATTTCTGGGACATCTGTCTAGNA	240
Db	181	ATGTATATCCTACAAAGTAGGATTTAGGAAGAGAAATTTCTGGGACATCTGTCTAGNA	240
QY	241	GGGTTTAAAAACATTTATGCGATTTGTGAGTGTGCGTGTGGAGAGTGTGTTTATATTTGTC	300
Db	241	GGGTTTAAAAACATTTATGCGATTTGTGAGTGTGCGTGTGGAGAGTGTGTTTATATTTGTC	300
QY	301	ACCGTATATGCGCAACATCCGGGGGCTTACTTCAGCTCCGGGCACTGTGGCG	353
Db	301	ACCGTATATGCGCAACATCCGGGGGCTTACTTCAGCTCCGGGCACTGTGGCG	353